

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 13:08:02 ; Search time 2.55496 Seconds
(without alignments)
3431.399 Million cell updates/sec

Title: US-09-914-286-3_COPY_3805_3886

Perfect score: 429
Sequence: 1 LACOTHEOQHTTLALVRSR.....RLPTTLAFDHPNPTTLTHL 82

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubpaa/PTC_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubpaa/PTCUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	253	59.0	3739	9 US-09-860-846-33	Sequence 33, Appl
2	253	59.0	3739	9 US-09-988-384B-33	Sequence 33, Appl
3	253	59.0	3739	9 US-09-836-821-33	Sequence 33, Appl
4	253	59.0	3739	9 US-09-793-708-2	Sequence 2, Appl
5	253	59.0	3739	10 US-09-861-289-33	Sequence 33, Appl
6	253	59.0	11877	9 US-09-860-846-6	Sequence 6, Appl
7	253	59.0	11877	9 US-09-836-821-6	Sequence 6, Appl
8	253	59.0	11877	10 US-09-861-289-6	Sequence 6, Appl
9	253	59.0	12199	9 US-09-988-384B-6	Sequence 6, Appl
10	240	55.9	3816	9 US-09-808-880-3	Sequence 3, Appl
11	236	55.0	4150	9 US-09-808-880-2	Sequence 3, Appl
12	234	54.5	1562	9 US-09-860-846-35	Sequence 35, Appl
13	234	54.5	1562	9 US-09-988-384B-35	Sequence 35, Appl
14	234	54.5	1562	9 US-09-836-821-35	Sequence 35, Appl
15	234	54.5	1562	9 US-09-793-708-3	Sequence 3, Appl
16	234	54.5	1562	10 US-09-861-289-35	Sequence 35, Appl
17	229	53.4	3519	9 US-09-808-880-4	Sequence 4, Appl
18	224	52.2	5215	9 US-09-860-846-2	Sequence 2, Appl
19	224	52.2	5215	9 US-09-988-384B-2	Sequence 2, Appl

20	224	52.2	5215	9 US-09-836-821-2	Sequence 2, Appl
21	224	52.2	5215	10 US-09-861-289-2	Sequence 2, Appl
22	209	48.7	4551	9 US-09-793-708-1	Sequence 1, Appl
23	209	48.7	4613	9 US-09-860-846-31	Sequence 31, Appl
24	209	48.7	4613	9 US-09-988-384B-31	Sequence 31, Appl
25	209	48.7	4613	9 US-09-836-821-31	Sequence 31, Appl
26	209	48.7	4613	10 US-09-861-289-31	Sequence 31, Appl
27	199	46.4	1616	9 US-09-712-363-262	Sequence 262, App
28	191	44.5	1346	9 US-09-793-708-4	Sequence 4, Appl
29	189	44.1	1346	9 US-09-860-846-37	Sequence 37, Appl
30	189	44.1	1346	9 US-09-988-384B-37	Sequence 37, Appl
31	189	44.1	1346	9 US-09-836-821-37	Sequence 37, Appl
32	189	44.1	1346	10 US-09-861-289-37	Sequence 37, Appl
33	178	41.5	1402	9 US-09-712-363-166	Sequence 166, App
34	164	38.2	3798	9 US-10-014-717-6	Sequence 6, Appl
35	164	38.2	7257	9 US-10-014-717-5	Sequence 5, Appl
36	111	25.9	352	10 US-09-073-009-14	Sequence 14, Appl
37	111	25.9	352	10 US-09-023-588-14	Sequence 14, Appl
38	105.5	24.6	1421	10 US-09-793-306-14	Sequence 14, Appl
39	101.5	23.7	1832	9 US-10-014-717-2	Sequence 2, Appl
40	101.5	23.7	1832	9 US-10-014-717-4	Sequence 4, Appl
41	93.5	21.8	2439	9 US-10-014-717-7	Sequence 7, Appl
42	90	21.0	1812	10 US-09-775-938A-38	Sequence 38, Appl
43	88	20.5	1827	9 US-09-712-363-261	Sequence 261, App
44	75.5	17.6	90	9 US-09-976-059-12	Sequence 12, Appl
45	67	15.6	517	9 US-09-738-626-4534	Sequence 4534, Ap

ALIGNMENTS

```
RESULT 1
US-09-860-846-33
Sequence 33, Application US/09860846
Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/860,846
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-860-846-33

Query Match      59.0%; Score 253; DB 9; Length 3739;
Best Local Similarity 62.2%; Pred. No. 3.6e-21;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

QY      1 LACOTHEOQHTTLALVRSRHTATVLTGTPDTPPDRAFDLGFDSLTAVELRNRLSRTT 60
DB      3560 LAAMTPDPRVAHNRDVRTHVATVLTGHTGTSRVDLEAFDGTGDSLTAVELRNRLNAAT 3619
QY      61 GLRLPTTLAFDHPNPTTLTHL 82
DB      3620 GLRLPTTLAFDHPNPTTLTHL 3641

RESULT 2
US-09-988-384B-33
Sequence 33, Application US/09988384B
Publication No. US20030073824A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
```

APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.536US1
CURRENT APPLICATION NUMBER: US/09/988.384B
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: PCT/US99/14398
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 53
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-988-384B-33

Query Match 59.0%; Score 253; DB 9; Length 3739;
Best Local Similarity 62.2%; Pred. No. 3.6e-21;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

QY 1 LAGOTHEOQHTTLALVRSHTATVLTGHTTPTTTPDRAFRDLGFDLSITAVELRNRLSRTT 60
DB 3560 LAAMTPDDRVAHLRDLVKTHTVATVLTGHTGTPSRVDLERAFRDTGFDLSITAVELRNRLNAAT 3619
QY 61 GLRLPTTLAFDHPNPTTLTHHL 82
DB 3620 GLRLPATLVFDHPTPGELAGHL 3641

RESULT 3
US-09-836-821-33
Sequence 33, Application US/09836821
Publication No. US20030087405A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/836.821
CURRENT FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-836-821-33

Query Match 59.0%; Score 253; DB 9; Length 3739;
Best Local Similarity 62.2%; Pred. No. 3.6e-21;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

QY 1 LAGOTHEOQHTTLALVRSHTATVLTGHTTPTTTPDRAFRDLGFDLSITAVELRNRLSRTT 60
DB 3560 LAAMTPDDRVAHLRDLVKTHTVATVLTGHTGTPSRVDLERAFRDTGFDLSITAVELRNRLNAAT 3619
QY 61 GLRLPTTLAFDHPNPTTLTHHL 82
DB 3620 GLRLPATLVFDHPTPGELAGHL 3641

RESULT 4
US-09-793-708-2
Sequence 2, Application US/09793708
Publication No. US20030104597A1
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary

APPLICANT: BETLIACH, Melanie C.
APPLICANT: BETLIACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002121
CURRENT APPLICATION NUMBER: US/09/793.708
CURRENT FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: US 09/657,440
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: US 09/320,878
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: US 09/141,908
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: US 09/073,538
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: US 08/846,247
PRIOR FILING DATE: 1997-04-30
PRIOR APPLICATION NUMBER: US 60/134,990
PRIOR FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-793-708-2

Query Match 59.0%; Score 253; DB 9; Length 3739;
Best Local Similarity 62.2%; Pred. No. 3.6e-21;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

QY 1 LAGOTHEOQHTTLALVRSHTATVLTGHTTPTTTPDRAFRDLGFDLSITAVELRNRLSRTT 60
DB 3560 LAAMTPDDRVAHLRDLVKTHTVATVLTGHTGTPSRVDLERAFRDTGFDLSITAVELRNRLNAAT 3619
QY 61 GLRLPTTLAFDHPNPTTLTHHL 82
DB 3620 GLRLPATLVFDHPTPGELAGHL 3641

RESULT 5
US-09-861-289-33
Sequence 33, Application US/09861289
Patent No. US20020110897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/861.289
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-861-289-33

Query Match 59.0%; Score 253; DB 10; Length 3739;
Best Local Similarity 62.2%; Pred. No. 3.6e-21;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

QY 1 LAGOTHEOQHTTLALVRSHTATVLTGHTTPTTTPDRAFRDLGFDLSITAVELRNRLSRTT 60
DB 3560 LAAMTPDDRVAHLRDLVKTHTVATVLTGHTGTPSRVDLERAFRDTGFDLSITAVELRNRLNAAT 3619
QY 61 GLRLPTTLAFDHPNPTTLTHHL 82

Db 3620 GLRLPATLVFDHPTGELAGHL 3641

```
RESULT 6
US-09-860-846-6
; Sequence 6, Application US/09860846
; Patent No. US2002016742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 11877
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-6
```

Query Match 59.0%; Score 253; DB 9; Length 11877;
Best Local Similarity 62.2%; Pred. No. 1.6e-20;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

Qy 1 LACQTHEOQHTTLALVRSIHATVAGHTTPTDIPPDRAFRDLGFDLSLTAVELRNRLSRTT 60
Db 8509 LAAMTPDDRAVHLRDVLRTHVATVLGHTGTSRVLDLRAFRDGTGDSLTAVELRNRLNAAT 8568
Qy 61 GLRLPTTLAFDHPNPPTLTTHHL 82
Db 8569 GLRLPATLVFDHPTGELAGHL 8590

```
RESULT 7
US-09-836-821-6
; Sequence 6, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 11877
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-836-821-6
```

Query Match 59.0%; Score 253; DB 9; Length 11877;
Best Local Similarity 62.2%; Pred. No. 1.6e-20;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

Qy 1 LACQTHEOQHTTLALVRSIHATVAGHTTPTDIPPDRAFRDLGFDLSLTAVELRNRLSRTT 60
Db 8509 LAAMTPDDRAVHLRDVLRTHVATVLGHTGTSRVLDLRAFRDGTGDSLTAVELRNRLNAAT 8568
Qy 61 GLRLPTTLAFDHPNPPTLTTHHL 82

Db 8569 GLRLPATLVFDHPTGELAGHL 8590

```
RESULT 8
US-09-861-289-6
; Sequence 6, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 11877
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-861-289-6
```

Query Match 59.0%; Score 253; DB 10; Length 11877;
Best Local Similarity 62.2%; Pred. No. 1.6e-20;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

Qy 1 LACQTHEOQHTTLALVRSIHATVAGHTTPTDIPPDRAFRDLGFDLSLTAVELRNRLSRTT 60
Db 8509 LAAMTPDDRAVHLRDVLRTHVATVLGHTGTSRVLDLRAFRDGTGDSLTAVELRNRLNAAT 8568
Qy 61 GLRLPTTLAFDHPNPPTLTTHHL 82
Db 8569 GLRLPATLVFDHPTGELAGHL 8590

```
RESULT 9
US-09-988-384B-6
; Sequence 6, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 6
; LENGTH: 12199
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-6
```

Query Match 59.0%; Score 253; DB 9; Length 12199;
Best Local Similarity 62.2%; Pred. No. 1.6e-20;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

Qy 1 LACQTHEOQHTTLALVRSIHATVAGHTTPTDIPPDRAFRDLGFDLSLTAVELRNRLSRTT 60
Db 8831 LAAMTPDDRAVHLRDVLRTHVATVLGHTGTSRVLDLRAFRDGTGDSLTAVELRNRLNAAT 8890
Qy 61 GLRLPTTLAFDHPNPPTLTTHHL 82

Db 8891 GLRLPALTVPDHPPTGELAGHL 8912

RESULT 10

US-09-808-880-3
; Sequence 3, Application US/09808880
; Publication No. US20030027287A1
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: McDaniel, Robert
; APPLICANT: Tany, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/808,880
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US/09/428,517
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/120,254
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/106,100
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3816
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-808-880-3

Query Match 55.9%; Score 240; DB 9; Length 3816;
Best Local Similarity 67.2%; Pred. No. 1,4e-19;
Matches 45; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Qy 16 LVNSHIVLGHPTTPDTPDPAFDLGSITAVELRNRLSRTTGLRPTTALFDPHPNP 75

Db 3640 LVNRHVALVIGHSGPEALDAQAFRDIGFDSLTAVELRNRLMATGRLRPTLVFDYFNP 3699

Qy 76 TLTTHL 82

Db 3700 SALADHL 3706

RESULT 11

US-09-808-880-2
; Sequence 2, Application US/09808880
; Publication No. US20030027287A1
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: McDaniel, Robert
; APPLICANT: Tany, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/808,880
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US/09/428,517
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/120,254
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/106,100
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant

; OTHER INFORMATION: Oleandolide PKS
US-09-808-880-2

Query Match 55.0%; Score 236; DB 9; Length 4150;
Best Local Similarity 57.3%; Pred. No. 4.6e-19;
Matches 47; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

Qy 1 LAGOTHOQHTTLAVRSHIATVLTGTTPTDTPDPAFDLGSITAVELRNRLSRTT 60

Db 2441 LAGITAVEGRRLVQLVREGAAVYLGHSGAAYAPDPAFDLGSITAVELRNRLTAT 2500

Qy 61 GLRLPTTALFDPHPPTTTHL 82

Db 2501 GLRLPALTVPDHPPTGELAGHL 2522

RESULT 12

US-09-860-846-35
; Sequence 35, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 1562
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-35

Query Match 54.5%; Score 234; DB 9; Length 1562;
Best Local Similarity 58.7%; Pred. No. 2.3e-19;
Matches 44; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

Qy 8 QOHTTLAVRSHIATVLTGTTPTDTPDPAFDLGSITAVELRNRLSRTTGLRPTT 67

Db 1400 ERPPALITLVRTAAVYLGHSSPDVAPGAFTELGFSITAVLNQSLSTVGANLPAT 1459

Qy 68 LAFDPHPPTTTHL 82

Db 1460 TVFDHPPTPALAHL 1474

RESULT 13

US-09-988-384B-35
; Sequence 35, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 35
; LENGTH: 1562
; TYPE: PRT

ORGANISM: Streptomyces venezuelae
US-09-988-384B-35

Query Match 54.5%; Score 234; DB 9; Length 1562;
Best Local Similarity 58.7%; Pred. No. 2.3e-19;
Matches 44; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 8 QOHTTLALVRSHTATVLTGHTTPTTIPDRAFRDLGFDLSLTAVELRNRLSRTTGLRLPTT 67
DB 1400 ERRPALTLVTRTHAAVAGHSPPRVAFGRAFTELGFDSLTAVOLRNQLSTVVGNRLLPAT 1459

QY 68 LAFDHPNPTTLTHHL 82
DB 1460 TVFDHPTPALAHL 1474

RESULT 14
US-09-836-821-35
Sequence 35, Application US/09836821
Publication No. US20030087405A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/836,821
CURRENT FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PasteSeq for Windows Version 3.0
SEQ ID NO 35
LENGTH: 1562
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-836-821-35

Query Match 54.5%; Score 234; DB 9; Length 1562;
Best Local Similarity 58.7%; Pred. No. 2.3e-19;
Matches 44; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 8 QOHTTLALVRSHTATVLTGHTTPTTIPDRAFRDLGFDLSLTAVELRNRLSRTTGLRLPTT 67
DB 1400 ERRPALTLVTRTHAAVAGHSPPRVAFGRAFTELGFDSLTAVOLRNQLSTVVGNRLLPAT 1459

QY 68 LAFDHPNPTTLTHHL 82
DB 1460 TVFDHPTPALAHL 1474

RESULT 15
US-09-793-708-3
Sequence 3, Application US/09793708
Publication No. US20030104597A1
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002121
CURRENT APPLICATION NUMBER: US/09/793,708
CURRENT FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: US 09/657,440
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: US 09/320,878
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: US 09/141,908
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: US 09/073,538

PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: US 08/846,247
PRIOR FILING DATE: 1997-04-30
PRIOR APPLICATION NUMBER: US 60/134,990
PRIOR FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1562
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-793-708-3

Query Match 54.5%; Score 234; DB 9; Length 1562;
Best Local Similarity 58.7%; Pred. No. 2.3e-19;
Matches 44; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 8 QOHTTLALVRSHTATVLTGHTTPTTIPDRAFRDLGFDLSLTAVELRNRLSRTTGLRLPTT 67
DB 1400 ERRPALTLVTRTHAAVAGHSPPRVAFGRAFTELGFDSLTAVOLRNQLSTVVGNRLLPAT 1459

QY 68 LAFDHPNPTTLTHHL 82
DB 1460 TVFDHPTPALAHL 1474

Search completed: June 17, 2003, 13:46:48
Job time : 3.55496 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:54:32 ; Search time 1.49039 Seconds
(without alignments)
1618.821 Million cell updates/sec

Title: US-09-914-286-3_COPY_3805_3886

Perfect score: 429
Sequence: 1 LAGQTHQHQHTTLAVRSH.....RLPTTLAFDHPNPTTLTHL 82

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA: *
2: /cgn2_6/ptodata/1/1aa/5A COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/5B COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6A COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/6B COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/BACKFILE1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	270	62.9	3729	2 US-08-804-227C-4	Sequence 4, Appl
2	256	59.7	3724	2 US-08-804-227C-10	Sequence 10, Appl
3	256	59.7	3724	2 US-08-804-198-4	Sequence 4, Appl
4	253	59.0	3739	3 US-09-320-878-2	Sequence 2, Appl
5	253	59.0	3739	4 US-09-105-537-3	Sequence 3, Appl
6	253	59.0	11877	4 US-09-105-537-6	Sequence 6, Appl
7	240	55.9	3816	4 US-09-428-517-3	Sequence 3, Appl
8	236	55.0	4150	4 US-09-428-517-2	Sequence 2, Appl
9	235	54.8	3567	2 US-07-642-734C-4	Sequence 4, Appl
10	235	54.8	3567	3 US-08-439-009A-4	Sequence 4, Appl
11	234	54.5	1562	3 US-09-320-878-3	Sequence 3, Appl
12	234	54.5	1562	4 US-09-105-537-35	Sequence 35, Appl
13	229	53.4	3519	4 US-09-428-517-4	Sequence 4, Appl
14	226	52.7	3170	4 US-09-036-987A-4	Sequence 4, Appl
15	226	52.7	3170	4 US-09-370-700-4	Sequence 4, Appl
16	224	52.0	5215	4 US-09-105-537-2	Sequence 2, Appl
17	223	52.0	1996	2 US-08-804-227C-9	Sequence 9, Appl
18	223	52.0	1996	2 US-08-804-158-3	Sequence 3, Appl
19	223	52.0	3491	2 US-07-642-734C-2	Sequence 2, Appl
20	223	52.0	3491	3 US-08-439-009A-2	Sequence 2, Appl
21	219	51.0	4472	2 US-08-804-227C-2	Sequence 2, Appl
22	217	50.6	4545	2 US-08-804-227C-14	Sequence 14, Appl
23	217	50.6	4550	2 US-08-804-227C-8	Sequence 8, Appl
24	217	50.6	4550	2 US-08-804-158-2	Sequence 2, Appl
25	216	50.3	4928	4 US-09-036-987A-5	Sequence 5, Appl
26	216	50.3	4928	4 US-09-370-700-5	Sequence 5, Appl
27	214	49.9	502	4 US-09-413-814-106	Sequence 106, App

28	213.5	49.8	1864	2 US-08-804-227C-3	Sequence 3, Appl
29	209	48.7	4551	3 US-09-320-878-1	Sequence 1, Appl
30	209	48.7	4613	4 US-09-105-537-31	Sequence 31, Appl
31	208	48.5	5588	4 US-09-036-987A-6	Sequence 6, Appl
32	208	48.5	5588	4 US-09-370-700-6	Sequence 6, Appl
33	207	48.3	2152	4 US-09-036-987A-3	Sequence 3, Appl
34	207	48.3	2152	4 US-09-370-700-3	Sequence 3, Appl
35	207	48.3	3170	2 US-07-642-734C-5	Sequence 5, Appl
36	207	48.3	3170	3 US-08-439-009A-5	Sequence 5, Appl
37	206.5	48.1	185	4 US-09-154-083-14	Sequence 14, Appl
38	203	47.3	1611	2 US-08-804-227C-5	Sequence 5, Appl
39	198	46.2	1891	2 US-08-804-227C-12	Sequence 12, Appl
40	198	46.2	1891	2 US-08-804-198-6	Sequence 6, Appl
41	191	44.5	1346	3 US-09-320-878-4	Sequence 4, Appl
42	189	44.1	1346	4 US-09-105-537-37	Sequence 37, Appl
43	185	43.1	2595	4 US-09-036-987A-2	Sequence 2, Appl
44	185	43.1	2595	4 US-09-370-700-2	Sequence 2, Appl
45	180	42.0	6095	4 US-09-144-085-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-08-804-227C-4
; Sequence 4, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kunstoss, Stuart A.
; APPLICANT: Roestock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-876-2459
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3729 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-804-227C-4
```

Query Match: 62.9%; Score 270; DB 2; Length 3729;
Best Local Similarity: 63.4%; Pred. No. 5.6e-25;
Matches: 52; Conservative: 9; Mismatches: 21; Indels: 0; Gaps: 0;

QY 1 LAGQTHQHQHTTLAVRSHIATVLTGTTEDTTPDRAFDLGPDSLTAVELNRSLRTT 60
DB 3547 LTGTLQEGALVAELRAAAVNLGHGSDSIPEDRAFDGLDLSITAVEMNRSLSAAT 3606
QY 61 GLRLPTTLAFDHPNPTTLTHL 82

Db 3607 GLRLPATLVFDHPTPGELAGHL 3628

RESULT 2

US-08-804-227C-10
Sequence 10, Application US/08804227C
Patent No. 5876991

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhstose, Stuart A.
APPLICANT: Rosteck, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Plant, Thomas G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 3724 amino acids
TYPE: amino acid
TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-804-227C-10

Query Match 59.7%; Score 256; DB 2; Length 3724;
Best Local Similarity 58.5%; Pred. No. 3.3e-23;

Matches 48; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

Db 1 LAGOTHEOQHTTLALVRSHTATVLTGHTTPTDTPDRAFDGFDSTLAVELNRRLSRTT 60

3548 LAGLSAERHRLGLVGEVAVALGHSAAEVRPDRPFREVGFDSTLAVELNRMAAVT 3607

QY 61 GLRLPTTLAFDHPNPTTLTHHL 82

Db 3608 GVRLPATLVFDHPTPALSSHL 3629

RESULT 3

US-08-804-198-4
Sequence 4, Application US/08804198
Patent No. 5945320

GENERAL INFORMATION:

APPLICANT: Burett, Stanley G.
APPLICANT: Kuhstose, Stuart A.
APPLICANT: Rao, Nagendra R.
APPLICANT: Richardson, Mark A.
APPLICANT: Rosteck, Paul R., Jr.
TITLE OF INVENTION: PLATENOILIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: USA

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/804,198
FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 3724 amino acids
TYPE: amino acid
TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-804-198-4

Query Match 59.7%; Score 256; DB 2; Length 3724;
Best Local Similarity 58.5%; Pred. No. 3.3e-23;

Matches 48; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

QY 1 LAGOTHEOQHTTLALVRSHTATVLTGHTTPTDTPDRAFDGFDSTLAVELNRRLSRTT 60

Db 3548 LAGLSAERHRLGLVGEVAVALGHSAAEVRPDRPFREVGFDSTLAVELNRMAAVT 3607

QY 61 GLRLPTTLAFDHPNPTTLTHHL 82

Db 3608 GVRLPATLVFDHPTPALSSHL 3629

RESULT 4

US-09-320-878-2
Sequence 2, Application US/09320878A
Patent No. 6117659

GENERAL INFORMATION:

APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert

TITLE OF INVENTION: RECOMBINANT NARONOLIDE POLYKETIDE SYNTHASE

FILE REFERENCE: 300622002120

CURRENT APPLICATION NUMBER: US/09/320,878A

EARLIER FILING DATE: 1999-05-27

EARLIER APPLICATION NUMBER: CIP OF 09/141,908

EARLIER FILING DATE: 1998-08-28

EARLIER APPLICATION NUMBER: CIP OF 09/073,538

EARLIER FILING DATE: 1998-05-06

EARLIER APPLICATION NUMBER: CIP OF 08/846,247

EARLIER FILING DATE: 1997-04-30

EARLIER APPLICATION NUMBER: 60/119,139

EARLIER FILING DATE: 1999-02-08

EARLIER APPLICATION NUMBER: 60/100,880

EARLIER FILING DATE: 1998-09-22

EARLIER APPLICATION NUMBER: 60/087,080

EARLIER FILING DATE: 1998-05-28

NUMBER OF SEQ ID NOS: 34

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 3739

TYPE: PRT

ORGANISM: Streptomyces venezuelae

US-09-320-878-2

Query Match 59.0%; Score 253; DB 3; Length 3739;
Best Local Similarity 62.2%; Pred. No. 8e-23;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

QY 1 LAGOTHEOQHTLLALVRSNIATVIGHTPTTIPDRAFRDLGFDLSITAVELRNRLSRTT 60
DB 3560 LAAMTPDDRVAHLVDLVRTHTVATVIGHTPSRVDLERAFRDTGFDLSITAVELRNRLNAAT 3619

QY 61 GLRLPTLAPDHPNPPTLTTHHL 82
DB 3620 GLRLPATLVDPDHPPTGELAGHL 3641

RESULT 5

US-09-105-537-33
Sequence 33, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438051
CURRENT APPLICATION NUMBER: US/09/105.537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatcSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-33

Query Match 59.0%; Score 253; DB 4; Length 3739;
Best Local Similarity 62.2%; Pred. No. 8e-23;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

QY 1 LAGOTHEOQHTLLALVRSNIATVIGHTPTTIPDRAFRDLGFDLSITAVELRNRLSRTT 60
DB 3560 LAAMTPDDRVAHLVDLVRTHTVATVIGHTPSRVDLERAFRDTGFDLSITAVELRNRLNAAT 3619

QY 61 GLRLPTLAPDHPNPPTLTTHHL 82
DB 3620 GLRLPATLVDPDHPPTGELAGHL 3641

RESULT 6

US-09-105-537-6
Sequence 6, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438051
CURRENT APPLICATION NUMBER: US/09/105.537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatcSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-6

Query Match 59.0%; Score 253; DB 4; Length 11877;
Best Local Similarity 62.2%; Pred. No. 3.8e-22;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

QY 1 LAGOTHEOQHTLLALVRSNIATVIGHTPTTIPDRAFRDLGFDLSITAVELRNRLSRTT 60
DB 8509 LAAMTPDDRVAHLVDLVRTHTVATVIGHTPSRVDLERAFRDTGFDLSITAVELRNRLNAAT 8568

QY 61 GLRLPTLAPDHPNPPTLTTHHL 82
DB 8569 GLRLPATLVDPDHPPTGELAGHL 8590

RESULT 7

US-09-428-517-3
Sequence 3, Application US/09428517
Patent No. 6251636
GENERAL INFORMATION:
APPLICANT: Belkacem, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/428.517
CURRENT FILING DATE: 1999-10-28
EARLIER APPLICATION NUMBER: 60/120,254
EARLIER FILING DATE: 1999-02-16
EARLIER APPLICATION NUMBER: 60/106,100
EARLIER FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 3816
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-428-517-3

Query Match 55.9%; Score 240; DB 4; Length 3816;
Best Local Similarity 67.2%; Pred. No. 3.7e-21;
Matches 45; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 16 LVRSNIATVIGHTPTTIPDRAFRDLGFDLSITAVELRNRLSRTTGLRLPTLAPDHPNP 75
DB 3640 LVRGHVALVYGHSGPEALIDAEQAFRDTGFDLSITAVELRNRLNAETGLRLPTLVDPYDNP 3699

QY 76 TTYLTHHL 82
DB 3700 SALADHL 3706

RESULT 8

US-09-428-517-2
Sequence 2, Application US/09428517
Patent No. 6251636
GENERAL INFORMATION:
APPLICANT: Belkacem, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/428.517
CURRENT FILING DATE: 1999-10-28
EARLIER APPLICATION NUMBER: 60/120,254
EARLIER FILING DATE: 1999-02-16
EARLIER APPLICATION NUMBER: 60/106,100
EARLIER FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4150
TYPE: PRT
ORGANISM: Artificial Sequence

Query Match 59.0%; Score 253; DB 4; Length 4150;
Best Local Similarity 62.2%; Pred. No. 3.8e-22;
Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: Oleandrolide PKS
US-09-428-517-2

Query Match 55.0%; Score 236; DB 4; Length 4150;
Best Local Similarity 57.3%; Pred. No. 1.3e-20;
Matches 47; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

QY 1 LAGQTHQOHTTLALVRSHTATVLTPTTPDPAFRDLGFDLSLTAVELNRRLSRTT 60
DB 2441 LAGLVAEQERLVLQVREQAVALGSGAVALPDRAPKDLGFDLSLTAVELNRRLTAT 2500

QY 61 GLRLPTLAPDHPNPTTLTHHL 82
DB 2501 GLRLPTLAPDHPNPTTLTHHL 2522

RESULT 9

US-07-642-734C-4
Sequence 4, Application US/07642734C
Patent No. 5824513
GENERAL INFORMATION:
APPLICANT: Katz, L
APPLICANT: Donadio, S
APPLICANT: Mcalpine, J B
TITLE OF INVENTION: Recombinant DNA Method for Producing
TITLE OF INVENTION: Erythromycin Analogs
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman
STREET: Abbott Laboratories D377/APed-2 One Abbott
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/642,734C
FILING DATE: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Danckerts, Andreas M
REGISTRATION NUMBER: 32652
REFERENCE/DOCKET NUMBER: 4952.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-642-734C-4

Query Match 54.8%; Score 235; DB 2; Length 3567;
Best Local Similarity 56.1%; Pred. No. 1.5e-20;
Matches 46; Conservative 10; Mismatches 26; Indels 0; Gaps 0;

QY 1 LAGQTHQOHTTLALVRSHTATVLTPTTPDPAFRDLGFDLSLTAVELNRRLSRTT 60
DB 3403 LAGRSSESDVAGLAELVRSHTAAVSGSADQLPERKAFDLGFDLSLAATELNRRLGTAT 3462

QY 61 GLRLPTLAPDHPNPTTLTHHL 82
DB 3463 GVRLPSTLVFDHPPLAVALHHL 3484

RESULT 10
US-08-439-009A-4
Sequence 4, Application US/08439009A
Patent No. 6004787
GENERAL INFORMATION:
APPLICANT: Donadio, S
APPLICANT: Katz, L
APPLICANT: Mcalpine, J B
TITLE OF INVENTION: Method of Directing Biosynthesis of
TITLE OF INVENTION: Specific Polypeptides
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven F. Weinstein
STREET: Abbott Laboratories D377/APed-2 One Abbott
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,009A
FILING DATE: 11-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 4952.US.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-439-009A-4

Query Match 54.8%; Score 235; DB 3; Length 3567;
Best Local Similarity 56.1%; Pred. No. 1.5e-20;
Matches 46; Conservative 10; Mismatches 26; Indels 0; Gaps 0;

QY 1 LAGQTHQOHTTLALVRSHTATVLTPTTPDPAFRDLGFDLSLTAVELNRRLSRTT 60
DB 3403 LAGRSSESDVAGLAELVRSHTAAVSGSADQLPERKAFDLGFDLSLAATELNRRLGTAT 3462

QY 61 GLRLPTLAPDHPNPTTLTHHL 82
DB 3463 GVRLPSTLVFDHPPLAVALHHL 3484

RESULT 11
US-09-320-878-3
Sequence 3, Application US/09320878A
Patent No. 6117659
GENERAL INFORMATION:
APPLICANT: Ashley, Gary
APPLICANT: Bertlach, Melanie C.
APPLICANT: Bertlach, Mary C.
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002120
CURRENT APPLICATION NUMBER: US/09/320,878A
FILING DATE: 1999-05-27
EARLIER APPLICATION NUMBER: CIP OF 09/141,908
EARLIER FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: CIP OF 09/073,538

Query Match 54.8%; Score 235; DB 3; Length 3567;
Best Local Similarity 56.1%; Pred. No. 1.5e-20;
Matches 46; Conservative 10; Mismatches 26; Indels 0; Gaps 0;

QY 1 LAGQTHQOHTTLALVRSHTATVLTPTTPDPAFRDLGFDLSLTAVELNRRLSRTT 60
DB 3403 LAGRSSESDVAGLAELVRSHTAAVSGSADQLPERKAFDLGFDLSLAATELNRRLGTAT 3462

QY 61 GLRLPTLAPDHPNPTTLTHHL 82
DB 3463 GVRLPSTLVFDHPPLAVALHHL 3484


```
/ EARLIER FILING DATE: 1998-05-06
/ EARLIER APPLICATION NUMBER: CIP OF 08/846,247
/ EARLIER FILING DATE: 1997-04-30
/ EARLIER APPLICATION NUMBER: 60/119,139
/ EARLIER FILING DATE: 1998-02-08
/ EARLIER APPLICATION NUMBER: 60/100,880
/ EARLIER FILING DATE: 1998-09-22
/ EARLIER APPLICATION NUMBER: 60/087,080
/ EARLIER FILING DATE: 1998-05-28
/ NUMBER OF SEQ ID NOS: 34
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 3
/ LENGTH: 1562
/ TYPE: PRT
/ ORGANISM: Streptomyces venezuelae
US-09-320-878-3

Query Match      54.5%; Score 234; DB 3; Length 1562;
Best Local Similarity 58.7%; Pred. No. 6.5e-21;
Matches 44; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY      8 OQHHTLLAVRSHIATVLTGHTTPTTIPDRAFRDGLGFDLSITAVELRNRLSRTTGLRLPTT 67
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1400 ERRPALTLVTRTHAAVLGHSSPDVAVGAFTELGLFDLSITAVOLRNQLSTVGNRLPAT 1459
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      68 LAFDHPNPTTLTHHL 82
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1460 TVFDHPTRPALAAHL 1474

RESULT 12
US-09-105-537-35
/ Sequence 35, Application US/09105537A
/ Patent No. 6265302
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, D.H.
/ APPLICANT: Liu, H.
/ APPLICANT: Xue, Y.
/ APPLICANT: Zhao, L.
/ TITLE OF INVENTION: DNA encoding methymycin and pikromycin
/ FILE REFERENCE: 600,438US1
/ CURRENT APPLICATION NUMBER: US/09/105,537A
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PasteSeq for Windows Version 3.0
/ SEQ ID NO 35
/ LENGTH: 1562
/ TYPE: PRT
/ ORGANISM: Streptomyces venezuelae
US-09-105-537-35

Query Match      54.5%; Score 234; DB 4; Length 1562;
Best Local Similarity 58.7%; Pred. No. 6.5e-21;
Matches 44; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY      8 OQHHTLLAVRSHIATVLTGHTTPTTIPDRAFRDGLGFDLSITAVELRNRLSRTTGLRLPTT 67
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1400 ERRPALTLVTRTHAAVLGHSSPDVAVGAFTELGLFDLSITAVOLRNQLSTVGNRLPAT 1459
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      68 LAFDHPNPTTLTHHL 82
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1460 TVFDHPTRPALAAHL 1474

RESULT 13
US-09-428-517-4
/ Sequence 4, Application US/09428517
/ Patent No. 6251636
/ GENERAL INFORMATION:
/ APPLICANT: Betlach, Mary C.
/ APPLICANT: Shan, Sanjay Kishan Kant
/ APPLICANT: McDaniel, Robert
/ APPLICANT: Tang, Li
```

```
/ TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
/ FILE REFERENCE: 30062-20029.00
/ CURRENT APPLICATION NUMBER: US/09/428,517
/ CURRENT FILING DATE: 1999-10-28
/ EARLIER APPLICATION NUMBER: 60/120,254
/ EARLIER FILING DATE: 1999-02-16
/ EARLIER APPLICATION NUMBER: 60/106,100
/ EARLIER FILING DATE: 1998-10-29
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 3519
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-428-517-4

Query Match      53.4%; Score 229; DB 4; Length 3519;
Best Local Similarity 53.7%; Pred. No. 8.3e-20;
Matches 44; Conservative 14; Mismatches 24; Indels 0; Gaps 0;

QY      1 LAGQTHEQHTLLAVRSHIATVLTGHTTPTTIPDRAFRDGLGFDLSITAVELRNRLSRTT 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1479 LSGLSRVQEEELVELVRAQAAVLVGHGSAQDVPAERAFELGLFDLSITAVELRNGLAAAT 1538
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      61 GLRLPTTLAPDHPNPTTLTHHL 82
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1539 GIRLPATMAPDHPNATATATATATATATATATATATATATATATATATATATATATAT 1560

RESULT 14
US-09-036-987A-4
/ Sequence 4, Application US/09036987A
/ Patent No. 6143526
/ GENERAL INFORMATION:
/ APPLICANT: Baltez, Richard H.
/ APPLICANT: Broughton, Marty C.
/ APPLICANT: Crawford, Kathryn P.
/ APPLICANT: Madduri, Krishnamurthy
/ APPLICANT: Merlo, Donald J.
/ APPLICANT: Treadway, Paul J.
/ APPLICANT: Turner, Jan R.
/ APPLICANT: Waldron, Clive
/ TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
/ TITLE OF INVENTION: Production
/ NUMBER OF SEQUENCES: 39
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Dow Agrosciences LLC Patent Department
/ STREET: 9330 Zionsville Road
/ CITY: Indianapolis
/ STATE: Indiana
/ COUNTRY: USA
/ ZIP: 46268
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/036,987A
/ FILING DATE: 09-MAR-1998
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Stuart, Donald R
/ REGISTRATION NUMBER: 28,479
/ REFERENCE/DOCKET NUMBER: 50,608
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (317)337-4816
/ TELEFAX: (317)337-4847
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
```

LENGTH: 3170 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-4

Query Match 52.7%; Score 226; DB 4; Length 3170;
Best Local Similarity 56.1%; Pred. No. 1.7e-19;
Matches 46; Conservative 11; Mismatches 25; Indels 0; Gaps 0;

QY 1 LAGQTHQOHTTLALVRSHTATVIGHTTPDTPDRAFRDGFDSLTAVELRNRLSRTT 60
DB 1458 LAELPEAQRREHLVDVCAQVATVIGHSREEVQPERAFRALGFSIMAVDLRNRLTTAT 1517

QY 61 GLRLPTTLAFDPNPPTTLTHHL 82
DB 1518 GLRLPTTVFDYPNPALAAHL 1539

RESULT 15

US-09-370-700-4
Sequence 4, Application US/09370700
Patent No. 6274350
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patti J
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 3170
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-4

Query Match 52.7%; Score 226; DB 4; Length 3170;
Best Local Similarity 56.1%; Pred. No. 1.7e-19;
Matches 46; Conservative 11; Mismatches 25; Indels 0; Gaps 0;

QY 1 LAGQTHQOHTTLALVRSHTATVIGHTTPDTPDRAFRDGFDSLTAVELRNRLSRTT 60
DB 1458 LAELPEAQRREHLVDVCAQVATVIGHSREEVQPERAFRALGFSIMAVDLRNRLTTAT 1517

QY 61 GLRLPTTLAFDPNPPTTLTHHL 82
DB 1518 GLRLPTTVFDYPNPALAAHL 1539

Search completed: June 17, 2003, 13:09:31
Job time: 1.49039 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:50:57 ; Search time 4.01696 Seconds
(without alignments)
2720.102 million cell updates/sec

Title: US-09-914-286-3_COPY_3805_3886

Perfect score: 429
Sequence: 1 LAGQTHQQHTLLALVRSH.....RLPTTLAFDHPNPTTLTHHL 82

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq.101002.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	429	100.0	3972	21	AA823749
2	429	100.0	3972	22	AA65264
3	429	100.0	3972	22	AA65268
4	429	100.0	5532	21	AA823752
5	429	100.0	5532	22	AA65267
6	270	62.9	3729	18	AAW22603
7	256	59.7	3724	18	AAW23708
8	256	59.7	3724	18	AAW22608
9	254	59.2	4881	21	AA823751
10	254	59.2	4881	22	AA65266

11	253	59.0	3739	21	AA818638	Amino acid sequenc
12	253	59.0	3739	21	AA777193	S. venezuelae macr
13	253	59.0	3739	21	AA777201	S. venezuelae plx
14	253	59.0	3739	21	AA777202	Narbolonide synthase
15	253	59.0	12139	21	AA777180	S. venezuelae plx
16	246	57.3	1721	19	AAW52847	A. mediterranei t1
17	245	57.1	6239	21	AA823750	S. avermectilis ave
18	245	57.1	6239	22	AA65265	Streptomyces averm
19	241	56.2	6797	22	AA831558	Pimaricin biosynth
20	240	55.9	3816	21	AA927208	S. antibioticus 8
21	239	55.7	1561	21	AA777202	S. venezuelae plx
22	237	55.2	1688	19	AAW52848	A. mediterranei t1
23	236	55.0	899	22	AA610153	Protein encoded by
24	236	55.0	4150	21	AA927207	S. antibioticus 8
25	236	55.0	11096	22	AA610129	Streptomyces nours
26	235	54.8	3567	14	AA844431	eryl region polype
27	235	54.8	9477	22	AA810144	Streptomyces nours
28	234	54.5	1560	21	AA767203	Narbolonide synthase
29	234	54.5	1562	21	AA818639	Amino acid sequenc
30	234	54.5	1562	21	AA777194	S. venezuelae macr
31	233	54.3	3562	22	AA822213	Polyketide synthase
32	233	54.3	5069	19	AAW52845	A. mediterranei t1
33	229	53.4	3519	21	AA927209	S. antibioticus 8
34	229	53.4	5435	22	AA610145	Streptomyces nours
35	226	52.7	3170	20	AA739299	Spnc a polyketide
36	226	52.7	3170	20	AA870967	S. spinosa protein
37	224	52.2	1891	19	AAW52844	Amycolatopsis med
38	224	52.2	3201	22	AA822214	Polyketide synthase
39	224	52.2	3413	19	AAW52849	A. mediterranei t1
40	224	52.2	4630	21	AA777177	S. venezuelae vep
41	223	52.0	1996	18	AAW22607	Platenolide synthase
42	223	52.0	1996	18	AAW23717	Platenolide synthase
43	223	52.0	3398	14	AA844430	eryl region polype
44	222	51.7	3192	22	AA810128	Streptomyces nours
45	220	51.3	3546	22	AA822212	Polyketide synthase

ALIGNMENTS

RESULT 1
AA823749
AA823749 standard; Protein; 3972 AA.
XX
AC AA823749;
XX
DT 10-JAN-2001 (first entry)
XX
DE S. avermectilis avermectin aglycon synthase protein SEQ ID NO:3.
XX
KW Streptomyces avermectilis; avermectin aglycon synthase; biosynthesis;
KW multifunctional enzyme; polyketide; avermectin; veterinary drug;
KW agrochemical.
XX
OS Streptomyces avermectilis.
XX
PN WO2000050605-A1.
XX
PD 31-AUG-2000.
XX
PF 23-FEB-2000; 2000WO-JP01041.
XX
PR 24-FEB-1999; 99JP-0046961.
XX
PA (KITA) KITASATO INST.
XX
PI Omura S, Ikeda H;
XX
DR WPI; 2000-565458/52.
XX
DR N-PSDB; AAA82301.
XX
PT Avermectin aglycone synthase DNA and proteins encoded by all or part of
it for the production of avermectin and its derivatives for drug and

PT agrochemical use -
 XX
 PS Claim 32; Page 203-223; 314pp; Japanese.
 XX
 CC The present sequence represents an avermectin aglycon synthase protein.
 CC Also described are: (1) polypeptides encoded by all or part of the DNA;
 CC (2) expression vectors containing the DNA; (3) host cells transformed by
 CC the vectors; (4) preparation of the polypeptides by culture of the
 CC transformants; (5) preparation of avermectin aglycon or its derivatives
 CC by culture of transformed avermectin-producing microorganisms; and (6)
 CC oligonucleotides of 5-60 bases in length containing sense or antisense
 CC sequences from the avermectin aglycon synthase DNA. The enzymes are
 CC useful for the production of modified forms of avermectin and of the
 CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
 CC and agrochemicals.
 XX
 SQ Sequence 3972 AA;
 Query Match 100.0%; Score 429; DB 21; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 2.3e-46;
 Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LAGQTHQOHTTLALVRSHTATVLTGHTTPTTPDPAFRLGFDLSITAVELNRLSRTT 60
 DB 3805 LAGQTHQOHTTLALVRSHTATVLTGHTTPTTPDPAFRLGFDLSITAVELNRLSRTT 3864
 QY 61 GLRLPTTLAFDHPNPTTLTHHL 82
 DB 3865 GLRLPTTLAFDHPNPTTLTHHL 3886
 RESULT 2
 ID AAG65264 standard; Protein: 3972 AA.
 XX AAG65264;
 AC
 DT 04-DEC-2001 (first entry)
 XX
 DE Streptomycetes avermectilis protein SEQ ID NO: 4.
 XX
 KM Avermectin aglycone synthase; AAS; avermectin derivative;
 KW drug production; veterinary drug; pesticide.
 XX
 OS Streptomycetes avermectilis.
 XX
 PN MO200162939-A1.
 PD 30-AUG-2001.
 XX
 PF 23-FEB-2001; 2001MO-JP01381.
 XX
 PR 24-FEB-2000; 2000JP-0047405.
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 PA (KITA) KITASATO INST.
 PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
 XX
 DR WPI, 2001-582053/65.
 DR N-PSDB; AAH79277.
 XX
 PT New modified avermectin aglycone synthase derived from Streptomycetes
 PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
 PT drugs and pesticides -
 XX
 PS Claim 4; Page 167-180; 257pp; Japanese.
 XX
 CC The present invention relates to the production of modified derivatives
 CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
 CC avermectilis. The activity of an acyl carrier protein (ACP),
 CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
 CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl

CC reductase (SR) and/or thioesterase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is an S. avermectilis protein.
 XX
 SQ Sequence 3972 AA;
 Query Match 100.0%; Score 429; DB 22; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 2.3e-46;
 Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LAGQTHQOHTTLALVRSHTATVLTGHTTPTTPDPAFRLGFDLSITAVELNRLSRTT 60
 DB 3805 LAGQTHQOHTTLALVRSHTATVLTGHTTPTTPDPAFRLGFDLSITAVELNRLSRTT 3864
 QY 61 GLRLPTTLAFDHPNPTTLTHHL 82
 DB 3865 GLRLPTTLAFDHPNPTTLTHHL 3886
 RESULT 3
 ID AAG65268 standard; Protein: 3972 AA.
 XX AAG65268;
 AC
 DT 04-DEC-2001 (first entry)
 XX
 DE Streptomycetes avermectilis protein derivative SEQ ID NO: 8.
 XX
 KM Avermectin aglycone synthase; AAS; avermectin derivative;
 KW drug production; veterinary drug; pesticide.
 XX
 OS Synthetic.
 XX
 PN MO200162939-A1.
 PD 30-AUG-2001.
 XX
 PF 23-FEB-2001; 2001MO-JP01381.
 XX
 PR 24-FEB-2000; 2000JP-0047405.
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 PA (KITA) KITASATO INST.
 PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
 XX
 DR WPI, 2001-582053/65.
 DR N-PSDB; AAH79279.
 XX
 PT New modified avermectin aglycone synthase derived from Streptomycetes
 PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
 PT drugs and pesticides -
 XX
 PS Claim 5; Page 235-248; 257pp; Japanese.
 XX
 CC The present invention relates to the production of modified derivatives
 CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
 CC avermectilis. The activity of an acyl carrier protein (ACP),
 CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
 CC reductase (BR) and/or thioesterase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is derived from an S.
 CC avermectilis protein.
 XX
 SQ Sequence 3972 AA;
 Query Match 100.0%; Score 429; DB 22; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 2.3e-46;
 Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LAGQTHQOHTTLALVRSHTATVLTGHTTPTTPDPAFRLGFDLSITAVELNRLSRTT 60

Db 3805 LACQTHQOHTTLALVRSHIATVLTGHTTPTDTPPPRAFRDLGFDLSLTAVELNRLSRTT 3864
 QY 61 GLRLPTTLAFDHPNPTTLTHHL 82
 Db 3865 GLRLPTTLAFDHPNPTTLTHHL 3886

RESULT 4

AAAB23752
 ID AAB23752 standard; Protein: 5532 AA.

AC AAB23752;

DT 10-JAN-2001 (first entry)

DE S. avermectilis avermectin aglycon synthase protein SEQ ID NO:6.

XX Streptomyces avermectilis; avermectin aglycon synthase; biosynthesis;

KW multifunctional enzyme; polyketide; avermectin; veterinary drug;

XX agrochemical.

OS Streptomyces avermectilis.

PN WO200050605-A1.

PD 31-AUG-2000.

PF 23-FEB-2000; 2000MO-JP01041.

PR 24-FEB-1999; 99JP-0046961.

PA (KITA) KITASATO INST.

PI Omura S, Ikeda H;

DR WPI; 2000-565458/52.

XX N-PSDB; AAA92302.

PT Avermectin aglycone synthase DNA and proteins encoded by all or part of

PT it for the production of avermectin and its derivatives for drug and

PS agrochemical use -

XX Claim 32; Page 278-305; 314pp; Japanese.

CC The present sequence represents an avermectin aglycon synthase protein.

CC Also described are: (1) polypeptides encoded by all or part of the DNA;

CC (2) expression vectors containing the DNA; (3) host cells transformed by

CC the vectors; (4) preparation of the polypeptides by culture of the

CC transformants; (5) preparation of avermectin aglycon or its derivatives

CC by culture of transformed avermectin-producing microorganisms; and (6)

CC oligonucleotides of 5-60 bases in length containing sense or antisense

CC sequences from the avermectin aglycon synthase DNA. The enzymes are

CC useful for the production of modified forms of avermectin and of the

CC intermediates in its biosynthesis, for use as drugs, veterinary drugs

CC and agrochemicals.

RESULT 5

AAAG65267
 ID AAG65267 standard; Protein: 5532 AA.

AC AAG65267;

DT 04-DEC-2001 (first entry)

DE Streptomyces avermectilis protein SEQ ID NO: 7.

XX Streptomyces avermectilis; avermectin derivative;

KW drug production; veterinary drug; pesticide.

OS Streptomyces avermectilis.

PN WO200162939-A1.

PD 30-AUG-2001.

PF 23-FEB-2001; 2001MO-JP01381.

PR 24-FEB-2000; 2000JP-0047405.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PI (KITA) KITASATO INST.

DR Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;

DR WPI; 2001-582053/65.

XX N-PSDB; AAH79278.

PT New modified avermectin aglycone synthase derived from Streptomyces

PT avermectilis used in production of 22,23-dihydroavermectin B1a used in

PT drugs and pesticides -

XX Claim 4; Page 217-235; 257pp; Japanese.

PS The present invention relates to the production of modified derivatives

CC of avermectin aglycone synthase (AAS) derived from Streptomyces

CC avermectilis. The activity of an acyl carrier protein (ACP),

CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),

CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl

CC reductase (ER) and/or thioesterase (TE) domain may be reduced or

CC suppressed. The process can be used in the production of drugs, veterinary

CC drugs and pesticides. The present sequence is an S. avermectilis protein.

XX Claim 4; Page 217-235; 257pp; Japanese.

XX The present sequence represents an avermectin aglycon synthase protein.

XX Also described are: (1) polypeptides encoded by all or part of the DNA;

XX (2) expression vectors containing the DNA; (3) host cells transformed by

XX the vectors; (4) preparation of the polypeptides by culture of the

XX transformants; (5) preparation of avermectin aglycon or its derivatives

XX by culture of transformed avermectin-producing microorganisms; and (6)

XX oligonucleotides of 5-60 bases in length containing sense or antisense

XX sequences from the avermectin aglycon synthase DNA. The enzymes are

XX useful for the production of modified forms of avermectin and of the

XX intermediates in its biosynthesis, for use as drugs, veterinary drugs

XX and agrochemicals.

Query Match 100.0%; Score 429; DB 22; Length 5532;
 Best Local Similarity 100.0%; Pred. No. 3,5e-46;
 Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6

AAAM22603
 ID AAM22603 standard; Protein: 3729 AA.

AC AAM22603;

DT 27-FEB-1998 (first entry)

DE Tyllactone synthase ORF3 protein.

XX Tyllactone synthase gene cluster; tyllG gene; multifunctional protein;

KW polyketide; tyllactone synthase; antibiotic; tyllactone;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

XX tyllactone synthase; tyllactone synthase; tyllactone synthase;

```

OS Streptomyces fradiae.
XX
XX Key Location/Qualifiers
FH Misc-difference 1
FT /note= "encoded by GTG"
FT 43..465
FT Domain /note= "ketosynthase domain, KS4"
FT 588..916
FT Domain /note= "acyltransferase domain, AT4"
FT 1178..1356
FT Domain /note= "ketoreductase domain, KR 4"
FT 1451..1534
FT Domain /note= "acyl carrier protein domain, ACP4"
FT 1556..1980
FT Domain /note= "ketosynthase domain, KS5"
FT 2095..2425
FT Domain /note= "acyltransferase domain, AT5"
FT 2451..2636
FT Domain /note= "dehydratase domain, DH2"
FT 3274..3453
FT Domain /note= "ketoreductase domain, KR5"
FT 3548..3631
FT Domain /note= "acyl carrier protein domain, ACP5"
XX
XX EP791655-A2.
XX 27-AUG-1997.
XX
XX 19-FEB-1997; 97EP-0301056.
XX
XX 22-FEB-1996; 96US-0012078.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Dehoff BS, Kuhstoss SA, Rosteck PR, Sutton KL;
XX WPI; 1997-418046/39.
XX N-PSDB; AAT80413.
XX
XX DNA encoding Streptomyces fradiae tyllactone synthase domain - for
XX production of tyllactone-related polyketide compounds
XX
XX Claim 9; Pages 86-98; 220pp; English.
XX
XX AAM22601-W22605 represent proteins encoded by the tyllactone synthase
XX gene cluster of the invention. The gene cluster is also referred to as
XX the tyll gene, and was isolated from Streptomyces fradiae. These
XX sequences are multifunctional proteins which direct the synthesis of the
XX polyketide tyllactone, isolated from Streptomyces fradiae. Tyllactone is
XX the basic building block of the antibiotic tyllisin. The DNA sequence can
XX be modified so as to alter the type of carboxylic acids incorporated
XX CC the number of carboxylic acids incorporated and/or the post-condensation
XX reactions performed, thereby resulting in novel tyllisin-related
XX polyketides.
XX
XX Sequence 3729 AA;
XX
XX Query Match 62.9%; Score 270; DB 18; Length 3729;
XX Best Local Similarity 63.4%; Pred. No. 1.6e-25;
XX Matches 52; Conservative 9; Mismatches 21; Indels 0; Gaps 0;
XX
XX 1 LAGQTHQOHTTLLATVRSHTATVLTGHTTDTTPPPDAPFDLGSDSLTAVELNRRLSRTT 60
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX LTLVLTQEGALVLAERAAVAALVGHGSDSIPEDRAFDLGSDSLTAVEMNRRLSAAAT 3606
XX 61 GLRLPTTLAFDHPNPTTLTHL 82
XX : : : : : : : : : : : : : : : : : : : : : :
XX GLRLPATLVVDHPTPGLAGHL 3628
XX
XX Db 3607 GLRLPATLVVDHPTPGLAGHL 3628
XX
XX RESULT 7
XX AAM23718
XX AAM23718 standard; Protein; 3724 AA.

```

```

XX
XX AAM23718;
XX AC
XX 27-FEB-1998 (first entry)
XX
XX Platenolide synthase ORF3 protein.
XX
XX Platenolide synthase gene cluster; platenolide production; srng gene;
XX multi-functional protein; macroide antibiotic; spiramycin.
XX
XX Streptomyces ambofaciens.
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "encoded by TGT"
FT 36..459
FT Domain /note= "ketosynthase domain, KS4"
FT 594..921
FT Domain /note= "acyltransferase domain, AT4"
FT 1177..1350
FT Domain /note= "ketosynthase domain, KS04"
FT 1459..1523
FT Domain /note= "acyl carrier protein domain, ACP4"
FT 1545..1969
FT Domain /note= "ketosynthase domain, KS5"
FT 2088..2424
FT Domain /note= "acyltransferase domain, AT5"
FT 2445..2619
FT Domain /note= "dehydratase domain, DH5"
FT 2912..3261
FT Domain /note= "enoylreductase domain, ER5"
FT 3266..3451
FT Domain /note= "ketoreductase domain, KR5"
FT 3551..3631
FT Domain /note= "acyl carrier protein domain, ACP5"
XX
XX EP791656-A2.
XX 27-AUG-1997.
XX
XX 19-FEB-1997; 97EP-0301066.
XX
XX 22-FEB-1996; 96US-0012050.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Burgett SG, Kuhstoss SA, Rao RN, Richardson MA;
XX Rosteck PR;
XX WPI; 1997-418047/39.
XX N-PSDB; AAT78508.
XX
XX DNA encoding Streptomyces ambofaciens platenolide synthase domain -
XX for production of spiramycin-related polyketide antibiotics
XX
XX Claim 8; Pages 54-66; 81pp; English.
XX
XX AAM2716-W23720 represent proteins encoded by the platenolide synthase
XX gene cluster of the invention. The gene cluster is also referred to as
XX the srng gene, and was isolated from Streptomyces ambofaciens. These
XX sequences are multi-functional proteins which direct the synthesis of
XX the polyketide platenolide. Platenolide is the basic building block of
XX the macroide antibiotic spiramycin. The DNA can be used to produce
XX compounds exhibiting antibiotic activity based on the platenolide
XX structure, including specifically the macroide antibiotic spiramycin and
XX CC spiramycin analogues and derivatives. Modifications of the platenolide
XX synthase DNA sequence can be made so as to change the number and type of
XX carboxylic acids incorporated into the growing polyketide chain and to
XX change the kind of post-condensation processing that is conducted.
XX
XX Sequence 3724 AA;
XX
XX Query Match 59.7%; Score 256; DB 18; Length 3724;

```

	Best Local Similarity	58.5%	Pred. No.	1.le-23;	
	Matches	48;	Conservative	12;	Mismatches 22; Indels 0; Gaps 0;
Oy	1	LACQTHEOQHITLLALVRSNIAIVGHTTTPDIPDRAPRDIGFDSLTVAVELRNRLSRTT	60		
Dd	3548	LAGLSAEERHRRLLGLVEGHVAVALVGHSAAEVRDPFRPEFGPDSLTVAVELRNMAAVT	3607		
Oy	61	GURLPPTIAPFDHPNPPTTLTHHL	82		
Db	3608	GVRLPATLVFDHPPTPALSLSHL	3629		
<hr/>					
RESULT 8					
AAM22608	ID	AAM22608 standard; Protein; 3724 AA.			
XX	AC	AAM22608;			
XX	DT	27-FEB-1998 (first entry)			
XX	DE	Platenolide synthase ORF3 protein.			
XX	KW	Tyactone synthase gene cluster; tylG gene; multifunctional protein,			
XX	KM	polyketide; tyactone synthesis; antibiotic; tylostin.			
XX	OS	Streptomyces ambofaciens.			
XX	FH	Key	Location/Qualifiers		
FT	Modified-site	1	/note= "encoded by TGT"		
FT	Domain	36..459	/note= "ketosynthase domain, KS4"		
FT	Domain	594..921	/note= "acyltransferase domain, AT4"		
FT	Domain	1177..1350	/note= "ketosynthase domain, KS04"		
FT	Domain	1459..1523	/note= "acyl carrier protein domain, ACP4"		
FT	Domain	1545..1969	/note= "ketosynthase domain, KS5"		
FT	Domain	2088..2424	/note= "acyltransferase domain, AT5"		
FT	Domain	2445..2619	/note= "dehydratase domain, DH5"		
FT	Domain	2912..3261	/note= "enoylreductase domain, ER5"		
FT	Domain	3266..3451	/note= "ketoreductase domain, KR5"		
FT	Domain	3551..3631	/note= "acyl carrier protein domain, ACP5"		
XX	PN	EPI791655-A2.			
XX	PD	27-AUG-1997.			
XX	PP	19-FEB-1997;	97EP-0301056.		
XX	PR	22-FEB-1996;	96US--0012078.		
XX	PA	(BLIL) LILLY & CO BLI.			
XX	PI	Dehoff BS, Kuhstoss SA, Rosteck PR, Sutton KL;			
XX	DR	MP1, 1997-418046/39.			
XX	N	N-P8DB; AAT80414.			
XX	PT	DNA encoding Streptomyces fradiae tyactone synthase domain - for			
XX	CC	production of tylostin-related polyketide compounds			
XX	BS	Example 2; Pages 155-167; 220pp; English.			
XX	AW	AAM22606-W22610 represent proteins encoded by the platenolide synthase			
XX	CC	gene cluster. The gene cluster is also referred to as the strg gene, and			

CC was isolated from Streptomyces ambofaciens. These sequences are
CC multi-functional proteins which direct the synthesis of the polyketide
CC platenolide. Platenolide is the basic building block of the macrocyclic
CC antibiotic spiramycin. The DNA encoding this sequence was used along
CC with the tylg gene (see AAT80413) to create a hybrid ORF1 sequence (see
CC AAT80415). The tylg gene is the tyactacon synthase gene cluster of the
CC invention. The tylg sequence was isolated from Streptomyces fradiae, and
CC encodes multifunctional proteins which direct the synthesis of the
CC polyketide tyactacon. Tyactacon is the basic building block of the
CC antibiotic tylosin. The hybrid sequence can be used to transform S.
CC ambofaciens lacking the smg ORF1 sequence, or S. fradiae lacking the
CC tylg ORF1 sequence, so that they can produce polyketides. The DNA
CC sequence can be modified so as to alter the type of carboxylic acids
CC incorporated, the number of carboxylic acids incorporated and/or the
CC post-condensation reactions performed, thereby resulting in novel
CC tylosin-related polyketides.

SO Sequence 3724 AA;

Query Match 59.7%; Score 256; DB 18; Length 3724;
Best Local Similarity 58.5%; Pred. No. 1,1e-23;
Matches 48; Conservative 12; Mismatches 22; Indels 0; Gaps 0;

Dy 1 LAGOTHEQOHTTLALVRSIHATVLTGTTEDTLPDRAPFDLGPDSTLAVELNRRLSRTT 60
Db 3548 LAGISAERHERDLRLGLGEHVAVALGHGSAAEVRPDRPREVGSPDLSLAVELNRMMAVT 3607

Dy 61 GLRLPTTLAFDHNPNTLTTHLL 82
Db 3608 GVRLPATLVVDHPRPALSSHL 3629

RESULT 9
AAB23751
ID AAB23751 standard; Protein; 4881 AA.

XX AAB23751;
XX
XX 10-JAN-2001 (first entry)
XX
XX DE S. avermitilis avermectin aglycon synthase protein SEQ ID NO:5.
XX
XX KM Streptomyces avermitilis; avermectin aglycon synthase; biosynthesis;
XX KM multifunctional enzyme; polyketide; avermectin; veterinary drug;
XX KM agrochemical.
XX
XX CS Streptomyces avermitilis.
XX
XX PN WO200050605-A1.
XX
XX PD 31-AUG-2000.
XX
XX PF 23-FEB-2000; 2000WO-JP01041.
XX
XX PR 24-FEB-1999; 99JP-0046961.
XX
XX PA (KITA) KITASATO INST.
XX
XX PI Omura S, Ikeda H;
XX WPI; 2000-565458/52.
XX DR N-PSTD; AAA92302.
XX
XX Avermectin aglycone synthase DNA and proteins encoded by all or part of
XX it for the production of avermectin and its derivatives for drug and
XX agrochemical use -
XX
XX Claim 32; Page 254-278; 314pp; Japanese.

The present sequence represents an avermectin aglycon synthase protein.
Also described are: (1) polypeptides encoded by all or part of the DNA;
(2) expression vectors containing the DNA; (3) host cells transformed by
the vectors; (4) preparation of the polypeptides by culture of the

transformants; (5) preparation of avermectin aglycon or its derivatives by culture of transformed avermectin-producing microorganisms; and (6) oligonucleotides of 5-60 bases in length containing sense or antisense sequences from the avermectin aglycone synthase DNA. The enzymes are useful for the production of modified forms of avermectin and of the intermediates in its biosynthesis, for use as drugs, veterinary drugs and agrochemicals.

CC Sequence 4881 AA;

Query Match 59.2%; Score 254; DB 21; Length 4881;
Best Local Similarity 62.2%; Pred. No. 2.9e-23;
Matches 51; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 1 LAGOTHEOQHTLLAVRSHIATVAGHTTPTTTPDPAFDDGFDSTLAVELNRRISRTT 60
DB 2525 LAGIDPAAQAEVTLVDLVTHAAVLGHTADAVPERAFDDGFDSTLAVELNRRINTAT 2584
QY 61 GLRLPTTLAFDHPNPPTTLTHHL 82
DB 2585 GLRFPRTLVFDHPRPVLAHAI 2606

RESULT 10
AAG65266
ID AAG65266 standard; Protein; 4881 AA.

XX AAG65266;

AC AAG65266;

DT 04-DEC-2001 (first entry)

DE Streptomyces avermiltillis protein SEQ ID NO: 6.

KM Avermectin aglycone synthase; AAS; avermectin derivative;

OS Streptomyces avermiltillis.

PN MO200162939-A1.

PD 30-AUG-2001.

PF 23-FEB-2001; 2001MO-JP01381.

PR 24-FEB-2000; 2000JP-0047405.

PA (KYOWA) KYOWA HAKKO KOGYO KK.

PA (KITA) KITASATO INST.

PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;

XX WPI; 2001-582053/65.

DR N-PSDB; AAH79278.

PT New modified avermectin aglycone synthase derived from Streptomyces
PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
PT drugs and pesticides -

PS Claim 4; Page 201-217; 257pp; Japanese.

CC The present invention relates to the production of modified derivatives
CC of avermectin aglycone synthase (AAS) derived from Streptomyces
CC avermiltillis. The activity of an acyl carrier protein (ACP),
CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC suppressed. The process can be used in the production of drugs, veterinary
CC drugs and pesticides. The present sequence is an S. avermiltillis protein.

XX Sequence 4881 AA;

Query Match 59.2%; Score 254; DB 22; Length 4881;
Best Local Similarity 62.2%; Pred. No. 2.9e-23;

Matches 51; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 1 LAGOTHEOQHTLLAVRSHIATVAGHTTPTTTPDPAFDDGFDSTLAVELNRRISRTT 60
DB 2525 LAGIDPAAQAEVTLVDLVTHAAVLGHTADAVPERAFDDGFDSTLAVELNRRINTAT 2584
QY 61 GLRLPTTLAFDHPNPPTTLTHHL 82
DB 2585 GLRFPRTLVFDHPRPVLAHAI 2606

RESULT 11
AAB18638
ID AAB18638 standard; Protein; 3739 AA.

XX AAB18638;

DT 22-JAN-2001 (first entry)

DE Amino acid sequence of narbonolide synthase subunit 2 (PICAI1).

KM Narbonolide synthase; polyketide synthase gene; narbonolide polyketide;

KM antibiotic; C12-hydroxylase; pick; desosamine biosynthesis;

KM desosaminyl transferase enzyme; ketolide; beta-glucosidase enzyme;

KM picromycin biosynthesis.

OS Streptomyces venezuelae.

XX US6117659-A.

PN 12-SEP-2000.

PF 27-MAY-1999; 99US-0320878.

PR 28-MAY-1998; 98US-0087080.

PR 22-SEP-1998; 98US-0100880.

PR 08-FEB-1999; 99US-0119139.

PR 20-MAY-1999; 99US-0134990.

PR 30-APR-1997; 97US-0846247.

PR 06-MAY-1998; 98US-0073538.

PR 28-AUG-1998; 98US-0141908.

PA (KOSA-) KOSAN BIOSCIENCES INC.

PI Ashley G, Betlach MC, Betlach M, Tang L, McDaniel R;

DR WPI; 2000-610844/58.

PT New recombinant pick hydroxylase gene of Streptomyces venezuelae useful
PT for converting ketolides to antibiotics and as antibiotics and
PT intermediates in the synthesis of compounds with pharmaceutical value

PS Disclosure; Columns 11-12; 117pp; English.

CC The present sequence represents a narbonolide synthase subunit 2
CC (PICAI1). The nucleotide sequence encoding it is used in the course of
CC the invention. The specification describes a recombinant DNA compound
CC expressing recombinant polyketide synthase genes in host cells for the
CC production of narbonolide, narbonolide derivatives and polyketides that
CC are useful as antibiotics and as intermediates in the synthesis of
CC compounds with pharmaceutical value. The DNA compounds may also encode
CC a C12-hydroxylase (pick), desosamine biosynthesis and desosaminyl
CC transferase enzymes (useful for conversion of ketolides to antibiotics),
CC and the beta-glucosidase enzyme (involved in picromycin biosynthesis).
CC These compounds are also useful for increasing the antibiotic activity
CC of a compound relative to the unhydroxylated compound. The recombinant
CC host cells are useful as genetic systems that allow rapid engineering
CC of the narbonolide polyketide synthase. These would be valuable for
CC creating novel ketolide analogs for pharmaceutical applications.

XX Sequence 3739 AA;

	Query Match	Similarity	59.0%	Score 253;	DB 21;	Length 3739;
	Best Local	Similarity	62.2%	Pred. NO. 2.8e-23;		
	Matches	51;	Conservative	7;	Mismatches	24;
					Indels	0;
					Gaps	0;
Dy	1	LACOTHEOCHITTLALVRSHTATVVGHTPTTIPEDRAFRDLGFDLSLTAVALRRNRSLRTT	60			
Db	3560	LAATPRDDRVANHLNDLVARTHTAATVGLGHOTPSRVDLERAFROTGRDLSLTAVELRRRLNAAT	3619			
Oy	61	GLRLPFTLAEPHNPFTLLTHNL	82			
Db	3620	GLRLPATLVFDPHPGELAGHL	3641			
RESULT 12						
AAAY7193	ID	AAAY7193 standard; Protein; 3739 AA.				
XX	AAAY7193;					
DT	05-JUN-2000	(first entry)				
DE						
XX						
XX						
KW		Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;				
KM		neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;				
KM		biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;				
KM		chronic obstructive pulmonary disease; respiratory inflammation;				
KM		hypercholesterolemia; crop protection agent.				
OS		Streptomyces venezuelae ATCC15439.				
PN		WO200000620-A2.				
XX						
PD	06-JAN-2000.					
PF	25-JUN-1999;	99WO-US14398.				
PR	26-JUN-1998;	98US-0105537.				
XX						
PA	(MINU) UNIV MINNESOTA.					
PI	Sherman DH, Liu H, Xue Y, Zhao L;					
DR	WPI, 2000-160679/14.					
N-PSDB	AAZ87298.					
PT	Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.					
synthesis of methymycin and pikromycin -						
Claim 19; Page 403-415; 438pp; English.						
XX						
CC	The invention relates to an isolated and purified nucleic acid segment					
CC	comprising a desosamine biosynthetic gene cluster, a fragment or its					
CC	biologically active variant, where the nucleic acid sequence is not					
CC	derived from the eryC gene cluster of Saccharopolyspora erythraea or					
CC	Streptomyces antibioticus. The invention also relates to a macrolide					
CC	biosynthetic gene cluster, or fragments thereof. The macrolide					
CC	biosynthetic gene cluster encodes proteins which synthesise methymycin,					
CC	pikromycin, neomethymycin, narbomycin or a combination of these					
CC	compounds. Recombinant or augmented cells comprising the desosamine					
CC	and/or macrolide biosynthetic gene clusters are useful for the					
CC	production of biologically active macrolides. The macrolide biosynthetic					
CC	proteins are useful for synthesis of methymycin, pikromycin,					
CC	neomethymycin and narbomycin. The alternative termination of polyketide					
CC	synthesis may be useful to prepare novel antibiotics and					
CC	polyhydroxyalkanoate (PHA) monomers. The compounds produced by the					
CC	recombinant host cells are useful as biopolymers, e.g., in packaging or					
CC	biomedical applications, to engineer PHA monomer syntheses or to prepare					
CC	biologically active agents, such as chemotherapeutics,					
CC	immunosuppressants, agents to treat asthma, chronic obstructive pulmonary					
CC	disease as well as other diseases involving respiratory inflammation,					
CC	cholesterol-lowering agents or macrolide-based antibiotics which are					
CC	active against a variety of organisms, e.g., bacteria, including					

CC	multi-drug resistant pneumococci and other respiratory pathogens, as well
CC	as viral parasitic pathogens, or as crop protection agents (e.g.,
CC	fungicides or insecticides) via expression of polyketides in plants.
CC	Sequences AA77190-77197 represent macrolide biosynthetic enzymes from
CC	Streptomyces venezuelae ATCC 15439, which are encoded by sequences
CC	AA28725-287302.
XX	
XX	
XX	
SQ	Sequence 3739 AA;
	Query Match 59.0%; Score 253; DB 21; Length 3739;
	Best Local Similarity 62.2%; Pred. No. 2.8e-23;
	Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;
Oy	1 LAGOTHEQHTTLALVRSHTATVLTGHTTPPTPPDPAFRDLGFDSTLAVELRNRLSRTT 60
Db	3560 LAATTPDPRVAMLDLVTHATVATVGHGTSRVDLERAFRDTGDSLTAVELRNRLNAT 3619
Oy	61 GLRLPTTLAPHPPTTLTHHL 82
Db	3620 GLRLPATLVFDPHPGSLAGHL 3641
RESULT 13	
ID	AA777201 standard; Protein: 3739 AA.
XX	
AC	AA777201;
XX	
DT	05-JUN-2000 (first entry)
DE	S. venezuelae pik (macrolide biosynthesis) gene cluster protein #2.
XX	
XX	Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
KW	neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
KW	biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;
KW	chronic obstructive pulmonary disease; respiratory inflammation;
KW	hypercholesterolaemia; crop protection agent.
XX	
OS	Streptomyces venezuelae ATCC15439.
XX	
PN	WO200000620-A2.
XX	
PD	06-JAN-2000.
XX	
PF	25-JUN-1999; 99WO-US14398.
XX	
PR	26-JUN-1998; 98US-0105537.
XX	
PA	(MINU) UNIV MINNESOTA.
XX	
PI	Sherman DH, Liu H, Xue Y, Zhao L;
DR	WPI: 2000-160679/14.
XX	
DR	N-PSDB; AA287318.
XX	
PT	Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.
XX	synthesis of methymycin and pikromycin -
XX	
PS	Disclosure; Figure 31, 438bp; English.
XX	
CC	The invention relates to an isolated and purified nucleic acid segment
CC	comprising a desosamine biosynthetic gene cluster, a fragment or its
CC	biologically active variant, where the nucleic acid sequence is not
CC	derived from the eryC gene cluster of Saccharopolyspora erythraea or
CC	Streptomyces antibioticus. The invention also relates to a macrolide
CC	biosynthetic gene cluster, or fragments thereof. The macrolide
CC	biosynthetic gene cluster encodes proteins which synthesise methymycin,
CC	pikromycin, neomethymycin, narbomycin or a combination of these
CC	compounds. Recombinant or augmented cells comprising the desosamine
CC	and/or macrolide biosynthetic gene clusters are useful for the
CC	production of biologically active macrolides. The macrolide biosynthetic
CC	proteins are useful for synthesis of methymycin, pikromycin,
CC	neomethymycin and narbomycin. The alternative termination of polyketide

CC synthesis may be useful to prepare novel antibiotics and
CC polyhydroxyalkanoate (PHA) monomers. The compounds produced by the
CC recombinant host cells are useful as biopolymers, e.g., in packaging or
CC biomedical applications, to engineer PHA monomer synthases or to prepare
CC biologically active agents, such as chemotherapeutics,
CC immunosuppressants, agents to treat asthma, chronic obstructive pulmonary
CC disease as well as other diseases involving respiratory inflammation,
CC cholesterol-lowering agents or macrolide-based antibiotics which are
CC active against a variety of organisms, e.g., bacteria, including
CC multi-drug resistant pneumococci and other respiratory pathogens, as well
CC as viral parasitic pathogens, or as crop protection agents (e.g.,
CC fungicides or insecticides) via expression of polyketides in plants.
CC Sequences AAY77200-Y77203 and AAY80997 represent proteins encoded by the
CC macrolide biosynthetic gene cluster (pik) from *Streptomyces venezuelae*
CC ATCC 15439, as given in figure 31.

XX
XX
SO Sequence 3739 AA;

Query Match 59.0%; Score 253; DB 21; Length 3739;
Best Local Similarity 62.2%; Pred. No. 2.8e-23;

Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

QY 1 LAGCTHEQQTTLALVRSHTATVLTGHTTPDTPPPAPRDGSDLTAVELRNRLSRTT 60

DB 3560 LAAMTPDDRVAHARDIVRTHVATVLTGHTPSRVDLERAFRDTGSDLTAVELRNRLNAAT 3619

QY 61 GLRLPTTLAFDHPNPPTLTTHLL 82

DB 3620 GLRLPATLVDPDHPPTGELAGHL 3641

RESULT 14

AAY67202 standard; protein; 3739 AA.

XX
XX
AC AAY67202;

DT 23-MAR-2000 (first entry)

XX Nardonolide synthase subunit 2 (PICAI1) protein sequence.

KW Nardonolide polyketide synthase; PKS; nardonolide synthase subunit 2;

XX PICAI1; antibiotic production; nardomycin; picromycin; ketolide.

OS *Streptomyces venezuelae*.

XX
XX
PN WO9961599-A2.

XX 02-DEC-1999.

PF 27-MAY-1999; 99WO-US11814.

XX 28-MAY-1998; 98US-0087080.

PR 28-AUG-1998; 98US-0141908.

PR 22-SEP-1998; 98US-0100880.

PR 08-FEB-1999; 99US-0119139.

XX
XX
PA (KOSA-) KOSAN BIOSCIENCES INC.

PI Ashley G, Betlach MC, Betlach M, McDaniel R, Tang L;

DR WPI; 2000-072618/06.

XX N-PSDB; AA256001.

XX
XX
PS Example 2; Page 13-14; 98pp; English.

CC This is the *Streptomyces venezuelae* nardonolide synthase subunit 2,
CC PICAI1 protein sequence. The invention relates to recombinant DNA
CC containing a coding sequence for a nardonolide polyketide synthase
(PKS). Polyketides are compounds synthesised from 2-carbon units through

CC a series of condensations and subsequent modifications. Modular PKSs are
CC responsible for the production of many antibiotics including picromycin.
CC The nardonolide PKS consists of a loading module, six extender modules,
CC and two thioester domains. Four proteins make up the nardonolide PKS
CC (PICAI, PICAI1, PICAI1 and PICAI1). PICAI includes the loading module
CC and extender modules 1 and 2, PICAI1 includes extender modules 3 and 4,
CC PICAI1 includes extender module 5 and PICAI1 includes extender module 6
CC and a type II thioesterase domain. The second type II thioesterase
CC domain is found on the PICB protein. The nucleotide sequences encoding
CC all of these proteins can be isolated in recombinant form from the
CC recombinant cosmid PKO5023-27 (see AA256001). Nardonolide is
CC desosaminylated in *S. venezuelae* to yield nardomycin, the desosamine
CC transferase enzyme is required for this conversion, and the desosamine
CC biosynthetic genes are also found in cosmid PKO5023-27. The recombinant
CC DNA of the invention is used to express, in transformed cells,
CC nardonolide (or its derivatives) or other ketolides (particularly
CC hybrids), which may then be converted (e.g. by other enzymes
CC recombinantly expressed in the same hosts) to polyketide antibiotics or
CC their intermediates. The antibiotics are useful in human or veterinary
CC medicine.

XX
XX
SO Sequence 3739 AA;

Query Match 59.0%; Score 253; DB 21; Length 3739;
Best Local Similarity 62.2%; Pred. No. 2.8e-23;

Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

QY 1 LAGCTHEQQTTLALVRSHTATVLTGHTTPDTPPPAPRDGSDLTAVELRNRLSRTT 60

DB 3560 LAAMTPDDRVAHARDIVRTHVATVLTGHTPSRVDLERAFRDTGSDLTAVELRNRLNAAT 3619

QY 61 GLRLPTTLAFDHPNPPTLTTHLL 82

DB 3620 GLRLPATLVDPDHPPTGELAGHL 3641

RESULT 15

AAY77180 standard; protein; 12199 AA.

XX
XX
AC AAY77180;

DT 05-JUN-2000 (first entry)

XX *S. venezuelae* pik gene cluster-encoded protein, SEQ ID NO:6.

KW Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;

KW neomethymycin; nardomycin; polyhydroxyalkanoate monomer synthase;

KW biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;

KW chronic obstructive pulmonary disease; respiratory inflammation;

KW hypercholesterolaemia; crop protection agent.

XX
XX
OS *Streptomyces venezuelae* ATCC15439.

XX WO200000620-A2.

XX 06-JAN-2000.

PF 25-JUN-1999; 99WO-US14398.

PR 26-JUN-1998; 98US-0105537.

XX (MINU) UNIV MINNESOTA.

PI Sherman DH, Liu H, Xue Y, Zhao L;

DR WPI; 2000-160679/14.

XX N-PSDB; AA287285.

XX
XX
PT Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.
XX Claim 19; Page 315-353; 438pp; English.

XX The invention relates to an isolated and purified nucleic acid segment
 CC comprising a desosamine biosynthetic gene cluster, a fragment or its
 CC biologically active variant, where the nucleic acid sequence is not
 CC derived from the eryC gene cluster of *Saccharopolyspora erythraea* or
 CC *Streptomyces antibioticus*. The invention also relates to a macrolide
 CC biosynthetic gene cluster, or fragments thereof. The macrolide
 CC biosynthetic gene cluster encodes proteins which synthesise methymycin,
 CC pikromycin, neomethymycin, narbomycin or a combination of these
 CC compounds. Recombinant or augmented cells comprising the desosamine
 CC and/or macrolide biosynthetic gene clusters are useful for the production
 CC of biologically active macrolides. The macrolide biosynthetic proteins
 CC are useful for synthesis of methymycin, pikromycin, neomethymycin and
 CC narbomycin. The alternative termination of polyketide synthesis may be
 CC useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA)
 CC monomers. The compounds produced by the recombinant host cells are useful
 CC as biopolymers, e.g., in packaging or biomedical applications, to
 CC engineer PHA monomer synthases or to prepare biologically active agents,
 CC such as chemotherapeutics, immunosuppressants, agents to treat asthma,
 CC chronic obstructive pulmonary disease as well as other diseases involving
 CC respiratory inflammation, cholesterol-lowering agents or macrolide-based
 CC antibiotics which are active against a variety of organisms, e.g.,
 CC bacteria, including multi-drug resistant pneumococci and other
 CC respiratory pathogens, as well as viral parasitic pathogens, or as crop
 CC protection agents (e.g., fungicides or insecticides) via expression of
 CC polyketides in plants. The present sequence represents a protein
 CC encoded by the macrolide biosynthetic gene cluster (pik) from
 CC *Streptomyces venezuelae* ATCC 15439.

XX Sequence 12199 AA;

Query Match 59.0%; Score 253; DB 21; Length 12199;

Best Local Similarity 62.2%; Pred. No. 1.3e-22;

Matches 51; Conservative 7; Mismatches 24; Indels 0; Gaps 0;

QY 1 LAGQTHQQTTLALVRSHTATVGGHTTPTTTPDRAFRDLGFDSTLAVELRNRLSRTT 60
 DB 8831 LAAMTPDDRAVNLHLDLVRTHTVATVIGHTGTPSRVDLBRAPFDTGFDSTLAVELRNRLNAAT 8890
 QY 61 GLRLPTTLAFDHPNPPTLLTHHL 82
 DB 8891 GLRLPATLVFDPHPTPGELAGHL 8912

Search completed: June 17, 2003, 13:02:35
 Job time : 5.01696 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:53:02 / Search time 3.97473 Seconds
(without alignments)
4458,180 Million cell updates/sec

Title: US-09-914-286-3_COPY_366_451
Perfect score: 424
Sequence: 1 LRLSPQQOQLDLVRAH.....RLPVTLLPDHTPAVAARL 86

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	424	100.0	3972	2	Q9S0R8 streptomyc
2	418	98.6	3626	2	Q9F779 streptomyc
3	214	50.5	6146	2	Q93H45 streptomyc
4	204	48.1	5435	2	Q9L4X2 streptomyc
5	200	47.2	3970	2	Q93H18 streptomyc
6	199	46.9	4613	2	Q9ZG15 streptomyc
7	199	46.9	4928	2	Q9ALM3 streptomyc
8	197	46.5	3524	2	Q93H86 streptomyc
9	197	46.5	5644	2	Q93H86 streptomyc
10	196	46.2	3654	2	Q30766 streptomyc
11	196	46.2	4150	2	Q9K1V4 streptomyc
12	195	46.0	1562	2	Q9ZG13 streptomyc
13	193	45.5	3352	2	Q93H83 streptomyc
14	193	45.5	3413	2	Q54593 amycolatops
15	193	45.5	4881	2	Q9S0R3 streptomyc
16	192	45.3	2152	2	Q9ALM5 streptomyc

17	190	44.8	3729	2	Q33956 streptomyc
18	190	44.8	6145	2	Q93H84 streptomyc
19	189	44.6	2100	2	Q30481 streptomyc
20	189	44.6	3576	16	Q9EX53 streptomyc
21	188	44.3	1191	2	Q9RHV0 streptomyc
22	188	44.3	6048	2	Q93H87 streptomyc
23	188	44.3	6315	2	Q9ADL6 streptomyc
24	186	43.9	3591	2	Q9K1E1 streptomyc
25	186	43.9	3613	2	Q93HJ1 streptomyc
26	185	43.6	3939	2	Q93HJ3 streptomyc
27	184.5	43.5	11096	2	Q9L4W3 streptomyc
28	184	43.4	1762	2	Q52546 streptomyc
29	184	43.4	1763	2	Q52790 streptomyc
30	184	43.4	5060	2	Q52545 streptomyc
31	184	43.4	5069	2	Q52789 streptomyc
32	184	43.4	8817	2	Q53840 streptomyc
33	183	43.2	4685	2	Q93HJ2 streptomyc
34	183	43.2	4735	2	Q54666 streptomyc
35	182	42.9	3739	2	Q9ZG14 streptomyc
36	182	42.9	7525	2	Q9K1E0 streptomyc
37	181	42.7	1762	2	Q30480 streptomyc
38	181	42.7	3170	2	Q9ALM4 streptomyc
39	180	42.5	6260	2	Q54299 streptomyc
40	180	42.5	9507	2	Q9EW11 streptomyc
41	180	42.5	10223	2	Q54296 streptomyc
42	179	42.2	1366	2	Q9L4W5 streptomyc
43	179	42.2	1412	2	Q93H88 streptomyc
44	179	42.2	1620	16	P96285 mycobacteri
45	179	42.2	6797	2	Q9X993 streptomyc

ALIGNMENTS

RESULT 1
Q9S0R8 PRELIMINARY; PRT; 3972 AA.
ID Q9S0R8
AC Q9S0R8; 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAY-2002 (TrEMBLrel. 20, Last annotation update)
DE Type I polyketide synthase AVE1.
GN AVE1.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Streptomyces.
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9380548; PubMed=10449723;
RA Ikeda H., Nonomiy T., Usami M., Ohta T., Omura S.;
RT "Organization of the biosynthetic gene cluster for the polyketide
antihelminthic macrocyclic avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL; AB032367; BAA84474.1; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantane-actin.
DR InterPro; IPR000834; Zn_cardoPept.
DR Pfam; PF00698; Acyl_transf. 3.
DR Pfam; PF00109; ketoacyl-synt. 2.
DR Pfam; PF02801; ketoacyl-synt. C' 2.
DR Pfam; PF00550; pp-binding. 3.
DR PROSITE; PS50075; ACP DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL SYNTHASE; 2.
DR PROSITE; PS00133; CARBOXYPEPT. ZN 2; UNKNOWN_2.
DR PROSITE; PS00012; PHOSPHOPANTETHEIN; 2.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 3972 AA; 416852 MW; 2A293695S032B1C3 CRC64;
Query Match 100.0%; Score 424; DB 2; Length 3972;
Best Local Similarity 100.0%; Pred. No. 4,4e-38;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LTAISPFOOQSLDLVRAHTMAVLNDGNERASDAGPSAFHGFDSVMGVELRNRL 60
DB 366 LTAISPFOOQSLDLVRAHTMAVLNDGNERASDAGPSAFHGFDSVMGVELRNRL 425
QY 61 SKATGLRPLVTLIFDHTTPAAVAARL 86
DB 426 SKATGLRPLVTLIFDHTTPAAVAARL 451

RESULT 2
Q9F7T9 PRELIMINARY; PRT; 3626 AA.
AC Q9F7T9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC31271;
RA Hong Y.-S., Lee J.U.;
RT "Targeted Gene Disruption of the avermectin O-methyltransferase gene
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275943; AAC09812.1;
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR003880; Pantane attach.
DR Pfam; PF00698; Acyl transfer; 3.
DR Pfam; PF00109; ketoacyl-synt; 2.
DR Pfam; PF02801; ketoacyl-synt_C; 2.
DR Pfam; PF00550; pp-binding; 2.
DR PROSITE; PS50075; ACP DOMAIN; 2.
DR PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN 1.
PT NON_TER 3626 AA; 380557 MW; 6272F5F088C1A8D0 CRC64;
SQ SEQUENCE 3626 AA; 380557 MW; 6272F5F088C1A8D0 CRC64;

Query Match 98.6%; Score 418; DB 2; Length 3626;
Best Local Similarity 98.8%; Pred. No. 1.8e-37;
Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LTAISPFOOQSLDLVRAHTMAVLNDGNERASDAGPSAFHGFDSVMGVELRNRL 60
DB 347 LTAISPFOOQSLDLVRAHTMAVLNDGNERASDAGPSAFHGFDSVMGVELRNRL 406
QY 61 SKATGLRPLVTLIFDHTTPAAVAARL 86
DB 407 SKATGLRPLVTLIFDHTTPAAVAARL 432

RESULT 3
Q93HU5 PRELIMINARY; PRT; 6146 AA.
AC Q93HU5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

```

```

RA Shinoe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakai Y., Hattori M.;
RT "Genomic sequence of an industrial microorganism Streptomyces
RT avermitilis: Deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; AB070940; BAB69192.1;
DR InterPro; IPR002106; AATRNA ligaseII.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR004410; FAD.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR000209; peptidase_S8.
DR InterPro; IPR003880; Pantane attach.
DR InterPro; IPR002364; OOR zeta crystal.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00698; Acyl transfer; 4.
DR Pfam; PF00107; adh_zinc; 1.
DR Pfam; PF00109; ketoacyl-synt; 4.
DR Pfam; PF02801; ketoacyl-synt_C; 4.
DR Pfam; PF00550; pp-binding; 4.
DR Pfam; PF00108; thiolase; 1.
DR TIGRfam; TIGR00128; fad; 4.
DR PROSITE; PS00339; AA TRNA LIGASE II_2; UNKNOWN 1.
DR PROSITE; PS50075; ACP DOMAIN; 4.
DR PROSITE; PS00606; B KETOACYL SYNTHASE; UNKNOWN 3.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN 1.
DR PROSITE; PS01152; OOR ZETA CRYSTAL; UNKNOWN 1.
DR PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN 1.
KW Phosphopantetheine.
SQ SEQUENCE 6146 AA; 637460 MW; 1D56F1108D458D00 CRC64;

Query Match 50.5%; Score 214; DB 2; Length 6146;
Best Local Similarity 53.5%; Pred. No. 2.4e-14;
Matches 46; Conservative 9; Mismatches 27; Indels 4; Gaps 1;

QY 1 LTAISPFOOQSLDLVRAHTMAVLNDGNERASDAGPSAFHGFDSVMGVELRNRL 60
DB 5980 LTAISFAERHRLVLDVREHTAGV-----GHGSADIDDDQAFKTIQFSLTVELRNRL 6035
QY 61 SKATGLRPLVTLIFDHTTPAAVAARL 86
DB 6036 RTATSLVPAVLVFDHTTPAALTAHL 6061

RESULT 4
Q9L4X2 PRELIMINARY; PRT; 5435 AA.
AC Q9L4X2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
OS Streptomyces noursei.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1971;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11455;
RA Brautaset T., Sekurova O.N., Stetter H., Ellingsen T.E., Strom A.R.,
RA Valla S., Zotchev S.B.;
RT "Biosynthesis of the polyene antifungal antibiotic nystatin in
RT Streptomyces noursei ATCC 11455: analysis of the gene cluster and
RT deduction of the biosynthetic pathway."
RL Chem. Biol. 7:395-403(2000).
DR EMBL; AF263912; AAF71767.1;
DR HSSP; P25715; IMLA.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002198; Adh_short.
DR InterPro; IPR002085; Adh_zn_family.

```

DR InterPro; IPR004410; FabD.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Ppantne_attach.
 DR Pfam; PF00698; Acyl_transf; 3.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt; 3.
 DR Pfam; PF02801; ketoacyl-synt_C; 3.
 DR Pfam; PF00550; pp-binding; 3.
 DR TIGRfam; TIGR00128; fabd; 3.
 DR PROSITE; PS50075; ACP_DOMAIN; 3.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 3.
 KW Phosphopantetheine; Transferase.
 SO SEQUENCE 5435 AA; 562659 MW; AA55465DF087A38C CRC64;

Query Match 48.1%; Score 204; DB 2; Length 5435;
 Best Local Similarity 55.2%; Pred. No. 2.8e-13;
 Matches 48; Conservative 6; Mismatches 27; Indels 6; Gaps 2;

Qy 1 LTLSPQQOOSLLDLVRAHTMAVLNDGNERRTASDAGPS-ASFAHGFPSVWGVELRNRL 59
 ID 093H18 PRELIMINARY; PRT; 3970 AA.
 AC Q93H18;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE Modular polyketide synthase.
 GN OLM4.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinoue M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: Deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL; AB070940; BAB69199.1; -
 DR InterPro; IPR001227; Ac_transferase.
 DR InterPro; IPR002085; Adh_zn_family.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Ppantne_attach.
 DR Pfam; PF00698; Acyl_transf; 2.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF02801; ketoacyl-synt_C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PS50075; ACP_DOMAIN; 2.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 DR PROSITE; PS00162; QOR_ZETA_CRYSTAL; UNKNOWN_1.
 KW Phosphopantetheine.
 SO SEQUENCE 3970 AA; 412657 MW; A93E583FAFA68C47 CRC64;

RESULT 5
 Q93H18 PRELIMINARY; PRT; 3970 AA.
 AC Q93H18;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE Modular polyketide synthase.
 GN OLM4.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinoue M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: Deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL; AB070940; BAB69199.1; -
 DR InterPro; IPR001227; Ac_transferase.
 DR InterPro; IPR002085; Adh_zn_family.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Ppantne_attach.
 DR Pfam; PF00698; Acyl_transf; 2.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF02801; ketoacyl-synt_C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PS50075; ACP_DOMAIN; 2.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 DR PROSITE; PS00162; QOR_ZETA_CRYSTAL; UNKNOWN_1.
 KW Phosphopantetheine.
 SO SEQUENCE 3970 AA; 412657 MW; A93E583FAFA68C47 CRC64;

Query Match 47.2%; Score 200; DB 2; Length 3970;

Best Local Similarity 51.2%; Pred. No. 5.3e-13;
 Matches 44; Conservative 9; Mismatches 29; Indels 4; Gaps 1;
 Qy 1 LTLSPQQOOSLLDLVRAHTMAVLNDGNERRTASDAGPS-ASFAHGFPSVWGVELRNRL 60
 ID 092G15 PRELIMINARY; PRT; 4613 AA.
 AC Q92G15;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Type I polyketide synthase pikA1.
 GN pikA1.
 OS Streptomyces venezuelae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=54571;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC15439;
 RX MEDLINE=9844533; PubMed=9770448;
 RX Xue Y., Zhao L., Liu H.W., Sherman D.H.;
 RA "A gene cluster for macroide antibiotic biosynthesis in streptomyces
 RT venezuelae: architecture of metabolic diversity.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12111-12116(1998).
 DR EMBL; AF079138; AAC69329.1; -
 DR HSSP; P25715; TMLA.
 DR InterPro; IPR001227; Ac_transferase.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Ppantne_attach.
 DR InterPro; IPR000169; Shprot_acetate.
 DR Pfam; PF00698; Acyl_transf; 3.
 DR Pfam; PF00109; ketoacyl-synt; 3.
 DR Pfam; PF02801; ketoacyl-synt_C; 3.
 DR Pfam; PF00550; pp-binding; 3.
 DR PROSITE; PS50075; ACP_DOMAIN; 3.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
 KW Phosphopantetheine; Transferase.
 SO SEQUENCE 4613 AA; 477281 MW; A3BAF8D37CEC9383 CRC64;

Query Match 46.9%; Score 199; DB 2; Length 4613;

Best Local Similarity 48.8%; Pred. No. 8.2e-13;
 Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

Qy 1 LTLSPQQOOSLLDLVRAHTMAVLNDGNERRTASDAGPS-ASFAHGFPSVWGVELRNRL 60
 ID 092G15 PRELIMINARY; PRT; 4613 AA.
 AC Q92G15;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Type I polyketide synthase pikA1.
 GN pikA1.
 OS Streptomyces venezuelae.
 SO SEQUENCE 4613 AA; 477281 MW; A3BAF8D37CEC9383 CRC64;

RESULT 7
 Q92G15 PRELIMINARY; PRT; 4613 AA.
 AC Q92G15;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Type I polyketide synthase pikA1.
 GN pikA1.
 OS Streptomyces venezuelae.
 SO SEQUENCE 4613 AA; 477281 MW; A3BAF8D37CEC9383 CRC64;

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae;
 OC Saccharopolyspora.
 OX NCBI_TaxID=60894;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21257765; PubMed=11358695;
 RA Madron C., Matsumura P., Roateck P.R. Jr., Broughton M.C.,
 RT Turner J., Madduri K., Crawford K.P., Merlo D.J., Baltz R.H.;
 "Cloning and analysis of the spinosad biosynthetic cluster of
 Saccharopolyspora spinosa."
 RL Chem. Biol. 8:487-499(2001).
 DR EMBL; AY007564; AAC3263.1; -.
 DR HSSP; P25715; IMMA.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002198; ADH short.
 DR InterPro; IPR004410; FAD.
 DR InterPro; IPR00794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantne attach.
 DR Pfam; PF00698; Acyl transf. 3.
 DR Pfam; PF00106; adh short; 1.
 DR Pfam; PF00109; ketoacyl-synt. 3.
 DR Pfam; PF02801; ketoacyl-synt_C; 3.
 DR Pfam; PF00550; pp-binding; 3.
 DR TIGRFAMs; TIGR00128; fadD; 3.
 DR PROSITE; PS00075; ACP DOMAIN; 3.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; 3.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 3.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 4928 AA; 517247 MW; C22C6361888F865A CRC64;

Query Match 46.9%; Score 199; DB 2; Length 4928;
 Best Local Similarity 50.0%; Pred. No. 8.9e-13;
 Matches 43; Conservative 13; Mismatches 26; Indels 4; Gaps 1;
 QY 1 LTALSPYQOOSLLDLYRAHTMAVLNDGERTASDAGPSAFHIGFDSVGVGLRNRL 60
 DB 4762 LAALPPEARHGHLLEVLAEFTASTLGHGDSAAVAQPDRTFAELGFDSTLAVELRNRL 4817
 QY 61 SKATGLRLPVTLLFDHTTPAAVARNL 86
 DB 4818 NAVTGLRLPPTLVDFDHTTPAAVARNL 4843

RESULT 8
 093H86 PRELIMINARY; PRT; 3524 AA.
 AC 093H86;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Modular polyketide synthase.
 GN pREB2.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinoue M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: Deducing the ability of producing secondary
 metabolites."
 RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RL EMBL; AB070949; BAB69304.1; -.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR001993; Mitoch_carrier.
 DR InterPro; IPR003880; Pantne attach.
 DR Pfam; PF00698; Acyl_transf. 2.

DR Pfam; PF00109; ketoacyl-synt. 2.
 DR Pfam; PF02801; ketoacyl-synt_C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PS00075; ACP DOMAIN; 2.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; UNKNOWN_2.
 DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_2.
 KM Phosphopantetheine.
 SQ SEQUENCE 3524 AA; 366619 MW; 6AA65759C461EBCC4 CRC64;

Query Match 46.5%; Score 197; DB 2; Length 3524;
 Best Local Similarity 52.9%; Pred. No. 1.8e-12;
 Matches 40; Conservative 16; Mismatches 26; Indels 4; Gaps 1;
 QY 1 LTALSPYQOOSLLDLYRAHTMAVLNDGERTASDAGPSAFHIGFDSVGVGLRNRL 60
 DB 3352 IAGLPSEEREAVLDVRSVAALV---GHSSAKTVQPDQAFGDSGLTAVELRNRL 3407
 QY 61 SKATGLRLPVTLLFDHTTPAAVARNL 86
 DB 3408 NQATGLRLPPTLVDFDHTTPAAVARNL 3433

RESULT 9
 093NX8 PRELIMINARY; PRT; 5644 AA.
 AC 093NX8;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE AmphJ.
 GN AMPHJ.
 OS Streptomyces nodosus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=40318;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Caffey P., Lynch S.V., Flood B.M., Finnan S.M., O'Leary M.;
 "The amphoterin biosynthetic gene cluster from Streptomyces
 nodosus." (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF357202; AAK73502.1; -.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh_zn_family.
 DR InterPro; IPR004410; FAD.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR Pfam; PF00698; Acyl transf. 3.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt. 3.
 DR Pfam; PF02801; ketoacyl-synt_C; 3.
 DR Pfam; PF00550; pp-binding; 3.
 DR TIGRFAMs; TIGR00128; fadD; 3.
 DR PROSITE; PS00075; ACP DOMAIN; 3.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; UNKNOWN_3.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_3.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_3.
 KM Phosphopantetheine.
 SQ SEQUENCE 5644 AA; 585775 MW; CFA35E3BA0D2658 CRC64;

Query Match 46.5%; Score 197; DB 2; Length 5644;
 Best Local Similarity 52.9%; Pred. No. 1.8e-12;
 Matches 46; Conservative 8; Mismatches 27; Indels 6; Gaps 2;
 QY 1 LTALSPYQOOSLLDLYRAHTMAVLNDGERTASDAGPSAFHIGFDSVGVGLRNRL 59
 DB 5474 LAALGAEGQHRTLVGLVROETAGVLSHGSG---AAVAADNDFFSLGFDSTLAVELRNRL 5528
 QY 60 LSAATGLRLPVTLLFDHTTPAAVARNL 86
 DB 5529 LSAATGLRLPPTLVDFDHTTPAAVARNL 5555


```

RESULT 10
ID 030766 PRELIMINARY; PRT; 3654 AA.
AC 030766;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polypeptide synthase modules 4 and 5.
GN NTD3.
OS Streptomyces caelestis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=36816;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL-2821;
RX MEDLINE=98053867; PubMed=9393718;
RA Kikavos S.J., Katz L., Steasi D.;
RT "Identification and characterization of the niddamycin polyketide
RT synthase genes from Streptomyces caelestis.";
RL J. Bacteriol. 179:7515-7522(1997).
DR EMBL; AF016585; AAC46026.1; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002085; Adh zn family.
DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR003880; Ppantne_attach.
DR InterPro; IPR002364; QOR_zeta_crystal.
DR Pfam; PF00698; Acyl transfer; 2.
DR Pfam; PF00107; adh_zinc; 1.
DR Pfam; PF00109; ketoacyl-synt; 2.
DR Pfam; PF02801; ketoacyl-synt_C; 2.
DR Pfam; PF00550; pp-binding; 2.
DR PROSITE; PS50075; ACP DOMAIN; 2.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_2.
DR PROSITE; PS01162; QOR_ZETA_CRYSTAL; 1.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 3654 AA; 379108 MW; 59825E49EDCD06FF CRC64;

Query Match 46.2%; Score 196; DB 2; Length 3654;
Best Local Similarity 52.3%; Pred. No. 1,de-12;
Matches 45; Conservative 9; Mismatches 28; Indels 4; Gaps 1;

OY 1 LTAISPQQOQLLDLVRAHTMAVLNDGNERTASDAGPSAFHGFDSVNGVELNRRL 60
DB 3471 LSLSDSERLDLTLELVELVAALVGLGSAVASD---PAREAGFDSLTAVELNRRL 3526
OY 61 SKATGLRLPVTLLFDHTTPAAVAARL 86
DB 3527 AATGRLRLPATLVFDHTPTALAGRL 3552

RESULT 11
ID 09KIV4 PRELIMINARY; PRT; 4150 AA.
AC 09KIV4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 8,8a-deoxyoleandolid synthase 1.
GN OLEA1.
OS Streptomyces antibioticus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1890;
RN
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=20363406; PubMed=10908114;
RA Shah S., Xue Q., Tang L., Carney J.R., Betlach M., McDaniel R.;
RT "Cloning, Characterization, and Heterologous Expression of a
RT Polyketide Synthase and P-450 Oxidase Involved in the Biosynthesis of
RT the Antibiotic Oleandomycin";
RL J. Antibiot. 53:502-508(2000).
DR EMBL; AF220951; AAF82408.1; -
DR HSSP; P25715; IMLA.
DR InterPro; IPR002106; AAcRNA_ligaseII.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR004410; Fald.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Ppantne_attach.
DR InterPro; IPR000169; SHprot_acetate.
DR Pfam; PF00698; Acyl transfer; 3.
DR Pfam; PF00109; ketoacyl-synt; 3.
DR Pfam; PF02801; ketoacyl-synt_C; 3.
DR Pfam; PF00550; pp-binding; 3.
DR TIGRPFAM; TIGR00128; fald; 3.
DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR PROSITE; PS50075; ACP DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 3.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 4150 AA; 435261 MW; 9383296C4C16647D CRC64;

Query Match 46.2%; Score 196; DB 2; Length 4150;
Best Local Similarity 50.0%; Pred. No. 1,de-12;
Matches 43; Conservative 10; Mismatches 29; Indels 4; Gaps 1;

OY 1 LTAISPQQOQLLDLVRAHTMAVLNDGNERTASDAGPSAFHGFDSVNGVELNRRL 60
DB 2441 LAGTVAEQRLLVQVREGQAVALVGLGSAEVA----PDRAFKDLGFSLSVVELNRRL 2496
OY 61 SKATGLRLPVTLLFDHTTPAAVAARL 86
DB 2497 NTAATGLRLPVTAVFDYAPALAGRL 2522

RESULT 12
ID 09ZG13 PRELIMINARY; PRT; 1562 AA.
AC 09ZG13;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Type I polyketide synthase PkAIII.
GN PkAIII.
OS Streptomyces venezuelae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=54571;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC15439;
RX MEDLINE=98445333; PubMed=9770448;
RA Xue Y., Zhao L., Liu H.W., Sherman D.H.;
RT "A gene cluster for macrolide antibiotic biosynthesis in streptomyces
RT venezuelae: architecture of metabolic diversity.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12111-12116(1998).
DR EMBL; AF079138; AAC69331.1; -
DR HSSP; P25715; IMLA.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Ppantne_attach.
DR Pfam; PF00698; Acyl transfer; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt_C; 1.
DR Pfam; PF00550; pp-binding; 1.
DR PROSITE; PS50075; ACP DOMAIN; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.

```

KW Phosphopantetheine; Transferase.
SQ SEQUENCE 1562 AA; 163593 MW; DBABA0DAAA15B63F CRC64;

Query March 46.0%; Score 195; DB 2; Length 1562;
Best Local Similarity 48.8%; Pred. No. 6.4e-13;
Matches 42; Conservative 11; Mismatches 29; Indels 4; Gaps 1;

QY 1 LTALSPTOOQSLLLDVRAHTMAVLNDGNERTASDAGPSASFALHGFDSVGVLELNRLL 60
DB 1393 ITALPEPERPALTLVTRTHAAVLGHSSPPRVA-----PGRAFTELGFDSLTAVQLNRLL 1448
QY 61 SKATGLRLPVTLLFDHTTPAAVAARL 86
DB 1449 STVVGNRLLPATVVDHPTPALAAHL 1474

RESULT 13

Q93H83 PRELIMINARY; PRT; 3352 AA.

AC Q93H83;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Modular polyketide synthase.
GN PTEA5.
OS Streptomyces avermiltillis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyocinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermiltillis: Deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL: AB070949; BAB69307.1; -
DR InterPro: IPR001227; Ac transferase.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR002362; Antenna_beta.
DR InterPro: IPR007944; Ketoacyl-synt.
DR InterPro: IPR003880; Ppantne attach.
DR InterPro: IPR001031; Thioesterase.
DR Pfam: PF00698; Acyl transfer; 2.
DR Pfam: PF00106; adh_short; 1.
DR Pfam: PF00109; ketoacyl-synt; 2.
DR Pfam: PF02801; ketoacyl-synt C; 2.
DR Pfam: PF00550; pp-binding; 2.
DR Pfam: PF00975; Thioesterase; 1.
DR PROSITE: PSS0075; ACP DOMAIN; 2.
DR PROSITE: PSS0069; ANTENNA_COMP_BETA; UNKNOWN_1.
DR PROSITE: PSS0060; B_KETOACYL_SYNTHASE; UNKNOWN_2.
DR PROSITE: PSS0012; PHOSPHOPANTHETHEINE; UNKNOWN_2.
KW Phosphopantetheine.
SQ SEQUENCE 3352 AA; 353395 MW; 66A5865BFD2027B7 CRC64;

Query Match 45.5%; Score 193; DB 2; Length 3352;
Best Local Similarity 50.0%; Pred. No. 2.7e-12;
Matches 43; Conservative 11; Mismatches 28; Indels 4; Gaps 1;

QY 1 LTALSPTOOQSLLLDVRAHTMAVLNDGNERTASDAGPSASFALHGFDSVGVLELNRLL 60
DB 2968 LALAPVEOEERLEHRAQSAVVL---GHASAEVLPTAHFLEGFDSLTAVQLNRLL 3023
QY 61 SKATGLRLPVTLLFDHTTPAAVAARL 86
DB 3024 SAATGLRLPGLAIFDHTTPARIKHL 3049

RESULT 14

O54593
ID O54593 PRELIMINARY; PRT; 3413 AA.

AC O54593;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Polyketide synthase.
GN RIFE.

OS Amycolatopsis mediterranei (Nocardia mediterranei).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
OX NCBI_TaxID=33910;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=5699;
RA Kim C.G., Yu T.W., Fryhle C., Handa S., Floss H.G.;
RL J. Biol. Chem. 0:0-0(1998).
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN=5699;
RA August P.R., Tang L., Yoon Y.J., Ning S., Mueller R., Yu T.W.,
RA Taylor M., Hoffmann D., Kim C.G., Zhang X., Hutchinson C.R.,
RA Floss H.G.;
RL Chem. Biol. 5:0-0(0002).
RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN=5699;
RA August P.R., Tang L., Yoon Y.J., Ning S., Mueller R., Hutchinson C.R.,
RA Floss H.G.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.

RC STRAIN=LBG A3136;
RA Schupp T., Toupet C., Engel N., Gott S.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF040570; AAC01714.1; -
DR EMBL: AJ223012; CAA11039.1; -
DR HSSP: P25715; IMLA.

DR InterPro: IPR001227; Ac transferase.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR003880; Ppantne attach.
DR Pfam: PF00698; Acyl transfer; 2.
DR Pfam: PF00109; ketoacyl-synt; 2.
DR Pfam: PF02801; ketoacyl-synt C; 2.
DR Pfam: PF00550; pp-binding; 2.
DR PROSITE: PSS0075; ACP DOMAIN; 2.
DR PROSITE: PSS0060; B_KETOACYL_SYNTHASE; 2.
DR PROSITE: PSS0012; PHOSPHOPANTHETHEINE; UNKNOWN_2.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 3413 AA; 353729 MW; FFA25AC4A3920AA4 CRC64;

Query Match 45.5%; Score 193; DB 2; Length 3413;
Best Local Similarity 47.7%; Pred. No. 2.7e-12;
Matches 41; Conservative 14; Mismatches 27; Indels 4; Gaps 1;

QY 1 LTALSPTOOQSLLLDVRAHTMAVLNDGNERTASDAGPSASFALHGFDSVGVLELNRLL 60
DB 1598 LAGLAABEGALLLDVKTQVALVIGHAGPEAVRAD---TAKDTGFDLSISVELNRLL 1653
QY 61 SKATGLRLPVTLLFDHTTPAAVAARL 86
DB 1654 REASGLKLPATLVFDVPTVALAARL 1679

RESULT 15

Q9S0R3 PRELIMINARY; PRT; 4881 AA.

AC Q9S0R3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 4.
GN AVEA4.

OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9380548; Pubmed=10449723;
 RA Ikeda H., Nonomiya T., Usami M., Ohta T., Omura S.;
 RT "Organization of the biosynthetic gene cluster for the polyketide
 anheimitic macroide avermectin in Streptomyces avermitilis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
 DR EMBL; AB032367; BAA8479.1; -.
 DR HSSP; P25715; IMLA.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR00794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantone attach.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00698; Acyl_transf; 3.
 DR Pfam; PF00109; ketoacyl-synt; 3.
 DR Pfam; PF02801; ketoacyl-synt_C; 3.
 DR Pfam; PF00550; pp-binding; 3-
 DR Pfam; PF00975; Thioesterase; 1.
 DR PROSITE; PS50075; ACP_DOMAIN; 3.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
 KW Phosphopantetheine; Transferase.
 SQ SEQUENCE 4881 AA; 510303 MW; F6568C1F01ACD37A CRC64;

Query Match 45.5%; Score 193; DB 2; Length 4881;
 Best Local Similarity 50.0%; Pred. No. 4.1e-12;
 Matches 43; Conservative 7; Mismatches 32; Indels 4; Gaps 1;

QY 1 LTALSPTQOSLLDLVRAHTMAVLNDGNERTASDAGPSASFAHGFDSVGVLELRRL 60
 Db 2525 LAGIDPAAQAEVTVLDLVLTAAAVL---CHGIDAVVPERAFRDLGFDLSITAVELRRL 2580
 QY 61 SKATGLRPVTLIFDHTTPAAVARRL 86
 Db 2581 NTATGLRFPRTLVPFDHPRPALAAHI 2606

Search completed: June 17, 2003, 13:07:52
 Job time : 4.97473 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:52:22 ; Search time 0.937857 Seconds

(without alignments)
3803.313 Million cell updates/sec

Title: US-09-914-286-3_COPY_366_451

Sequence: 1 LTRALSPQQGSLLDLVRAH.....RLPVTLLFDHTTPAAVAVARL 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query #	Match Length	DB ID	Description
1	174	41.0	3567	1	ERY2_SACER
2	173	40.8	3491	1	ERY1_SACER
3	166	39.2	3519	1	OL56_STRAT
4	138	32.5	3172	1	ERY3_SACER
5	99	23.3	1538	1	PPS8_MYCTU
6	90.5	21.3	1774	1	MSAS_PENPA
7	84.5	19.9	4427	1	PKSL_BACSU
8	81	19.1	1876	1	PPSA_MYCTU
9	81	19.1	1876	1	PPSA_MYCTU
10	78.5	18.5	1278	1	DHBF_BACSU
11	78	18.4	838	1	PAS_MOUSE
12	73.5	17.3	2511	1	PAS_CHICK
13	72.5	17.1	331	1	GALE_LACCA
14	72	17.0	740	1	PURL_LACCA
15	72	17.0	1392	1	LYS2_YEAST
16	72	17.0	3587	1	SRP1_BACSU
17	70	16.5	2560	1	PPS2_BACSU
18	68.5	16.2	3587	1	SRP2_BACSU
19	68	16.0	4273	1	PKSM_BACSU
20	66	15.6	6486	1	TYCC_BACRA
21	65.5	15.4	398	1	BMPB_XENLA
22	65	15.3	362	1	SDHL_RAT
23	64.5	15.2	64	1	PAS_RABT
24	64	15.1	1419	1	LYS2_SCHPO
25	64	15.1	2110	1	MCAS_MYCBO
26	63.5	15.0	247	1	ATPI_MESVI
27	63.5	15.0	332	1	GALE_STRTR
28	63.5	15.0	454	1	CBPH_LOPAM
29	63	14.9	887	1	ACOC_CAREL
30	62.5	14.7	553	1	PUR3_SNYX3
31	62	14.6	87	1	ACPY_STRAM
32	62	14.6	150	1	RBFA_STRCO
33	61.5	14.5	741	1	LOXB_PHAUV

34	61	14.4	178	1	IF3_UREPA	Q9pqr4 ureaplasma
35	61	14.4	636	1	GIDA_BRUM	Q8yja5 bruceella me
36	61	14.4	1082	1	A3B2_HUMAN	Q13367 home sapien
37	61	14.4	3649	1	ACVS_NOCIA	P27743 nocardiia la
38	60.5	14.3	247	1	ATPI_MAIZE	P17244 zea mays (m
39	60.5	14.3	247	1	ATPI_ORYSA	P12083 oryza sativ
40	60.5	14.3	247	1	ATPI_WHEAT	Q9xpco triticum ae
41	60.5	14.3	415	1	LOLC_NEIMA	P57061 neisseria m
42	60.5	14.3	415	1	LOLC_NEIMA	P57062 neisseria m
43	60.5	14.3	530	1	MURE_MYCLE	O69557 mycobacteri
44	60.5	14.3	692	1	GLND_CORGL	Q9x706 corynebacte
45	60	14.2	213	1	HIS5_CAUCR	Q9a231 caulobacter

ALIGNMENTS

RESULT 1
ERY2_SACER STANDARD; PRT; 3567 AA.
AC 003132; 054096;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Erythronolide synthase, modules 3 and 4 (EC 2.3.1.94) (ORF 2) (6-deoxyerythronolide B synthase II) (DEBS 2).
GN ERYA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Actinobacteria; Actinobacteridae (class); Actinobacteridae; Actinomycetales; Pseudonocardiales; Pseudonocardaceae;
OC Saccharopolyspora.
OX NCBI_Taxid=1836;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91220065; PubMed=2024119;
RA Donadio S., Steaver M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT "Modular organization of genes required for complex polyketide biosynthesis.";
RL Science 252:675-679 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RX MEDLINE=92155230; PubMed=1740151;
RA Beville D.J., Cortes J., Haydock S.F., Leadley P.F.;
RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme.";
RT Eur. J. Biochem. 204:39-49 (1992).
RL -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-deoxyerythronolide B.
CC -1- COFACTOR: NADP. CONTRAINS 2 COVALENTLY BOUND PHOSPHOTRANSFERINES.
CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THIS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR) DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.
CC -1- SIMILARITY: TO PATTY ACID SYNTHASE (PAS).
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC -----
 DR EMBL; M63677; AAA26494.1; -;
 DR EMBL; X62569; CAA44448.1; -;
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Ach zn family.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Pantene_attach.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl_transf; 2.
 DR Pfam; PF02801; ketoacyl-synt C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
 DR PROSITE; PS00066; B KETOACYL SYNTHASE; 2.
 DR PROSITE; PS00075; ACP DOMAIN; 2.
 DR TRANSFERASE; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KW Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1484
 FT DOMAIN 1485 3567 MODULE 3.
 FT DOMAIN 27 488 MODULE 4.
 FT DOMAIN 559 884 BETA-KETOACYL SYNTHASE 1.
 FT DOMAIN 1130 1301 BETA-KETOACYL REDUCTASE 1 (POSSIBLY
 FT NON-FUNCTIONAL).
 FT DOMAIN 1397 1467 ACYL CARRIER (ACP) 1.
 FT DOMAIN 1485 1943 BETA-KETOACYL SYNTHASE 2.
 FT DOMAIN 2013 2336 ACYLTRANSFERASE (AT) 2.
 FT DOMAIN 2383 3066 DEHYDRATASE/ENOYLREDUCTASE (DH/ER).
 FT DOMAIN 3139 3322 BETA-KETOACYL REDUCTASE 2.
 FT DOMAIN 3415 3485 ACYL CARRIER (ACP) 2.
 FT ACT SITE 202 202 THIOESTER BOND.
 FT ACT SITE 651 651 ACYL-ENZYME INTERMEDIATE.
 FT BINDING 1430 1430 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT BINDING 1661 1661 THIOESTER BOND.
 FT ACT SITE 1661 1661 ACYL-ENZYME INTERMEDIATE.
 FT ACT SITE 2115 2115 NADP (ER).
 FT NP_BIND 2961 2978 NADP (KR).
 FT NP_BIND 3142 3157 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT BINDING 3448 3448 R -> A (IN REF. 2).
 FT CONFLICT 438 438 T -> S (IN REF. 2).
 FT CONFLICT 480 480 L -> F (IN REF. 2).
 FT CONFLICT 1241 1241 L -> V (IN REF. 2).
 FT CONFLICT 2664 2664 G -> V (IN REF. 2).
 SQ SEQUENCE 3567 AA; 374413 MW; E6284F4738A0C0 CRC64;
 Query Match 41.08; Score 174; DB 1; Length 3567;
 Best Local Similarity 48.98; Pred. No. 1.7e-11;
 Matches 44; Conservative 7; Mismatches 27; Indels 12; Gaps 2;
 QY 1 LFTALPTQOQLLDLVRAHTMAVLN---DDGNERTASDAGPSAFHLGPDVSGVGL 56
 DB 3403 LAGRSBDQVAGLAELVRSIAAAYSGVGSADQPERKA-----FKDLGFSIAVEL 3454
 QY 57 RNRLSKATGRLPVTLLFDHTTPAAVAARL 86
 DB 3455 RNRIGTATGRLPSTLVFDHPTPLAAVEHL 3484
 RESULT 2
 ERYL SAKER STANDARD; PRT; 3491 AA.
 AC 003131;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Erythronolide synthase, modules 1 and 2 (BC 2.3.1.94) (ORF 1) (6-
 DE deoxyerythronolide B synthase I) (DEBS 1).
 GN ERYA.
 OS Saccharopolyspora erythraea (Streptomyces erythraeus).
 OC Bacteria; Actinobacteria; Actinobacteriia (class); Actinobacteridae;
 OC Actinomycetales; Pseudonocardiales; Pseudonocardaceae;
 OC Saccharopolyspora.

OX NCBI_TaxID=1836;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220065; PubMed=2024119;
 RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
 RT "Modular organization of genes required for complex polyketide
 RT biosynthesis.";
 RL Science 252:675-679 (1991).
 RN [2]
 RP SEQUENCE OF 3474-3491 FROM N.A.
 RX MEDLINE=93231529; PubMed=8386127;
 RA Donadio S., Staver M.J.;
 RT "IS1136, an insertion element in the erythromycin gene cluster of
 RT Saccharopolyspora erythraea.";
 RL Gene 126:147-151 (1993).
 CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
 CC deoxyerythronolide B.
 CC -1- COFACTOR: NADP; CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTHETHEINES.
 CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
 CC BIOSYNTHESIS.
 CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH
 CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THIS ERYA SHOWING 3
 CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
 CC SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS
 CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
 CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
 CC RESPECTIVELY.
 CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),
 CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER
 CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
 CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
 CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
 CC OF THE FULL-LENGTH CHAIN.
 CC -1- SIMILARITY: NO PARTIAL ACID SYNTHASE (PAS).
 CC -1- SIMILARITY: CONTAINS 3 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC -----
 DR EMBL; M63676; AAA26493.2; -;
 DR EMBL; L07626; L07626; AAA26504.1; -;
 DR HSSP; P25715; TMUA.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Pantene_attach.
 DR Pfam; PF00106; ach_short; 1.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF00550; pp-binding; 3.
 DR Pfam; PF00698; Acyl_transf; 3.
 DR Pfam; PF02801; ketoacyl-synt C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 3.
 DR PROSITE; PS00066; B KETOACYL SYNTHASE; 2.
 DR PROSITE; PS00075; ACP DOMAIN; 3.
 DR TRANSFERASE; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KW Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1972
 FT DOMAIN 1979 3491 MODULE 1.
 FT DOMAIN 1 375 MODULE 2.
 FT DOMAIN 414 484 ACYLTRANSFERASE (AT) 1.
 FT DOMAIN 503 961 ACYL CARRIER (ACP) 1.
 FT DOMAIN 1030 1356 BETA-KETOACYL SYNTHASE 1.
 FT DOMAIN 1611 1794 ACYLTRANSFERASE (AT) 2.
 FT DOMAIN 1888 1958 BETA-KETOACYL REDUCTASE 1.
 FT DOMAIN 1979 2441 ACYL CARRIER (ACP) 2.
 FT DOMAIN 2507 2854 BETA-KETOACYL SYNTHASE 2.
 FT DOMAIN 3055 3237 ACYLTRANSFERASE (AT) 3.
 FT DOMAIN 3334 3404 BETA-KETOACYL REDUCTASE 2.
 FT DOMAIN 3334 3404 ACYL CARRIER (ACP) 3.

FT	ACT_SITE	145	145	ACYL-ENZYME INTERMEDIATE.
FT	BINDING	447	447	PHOSPHOPANTHETHEINE (BY SIMILARITY).
FT	ACT_SITE	677	677	BETA-KETOACYL SYNTHASE.
FT	ACT_SITE	1128	1128	ACYL-ENZYME INTERMEDIATE.
FT	NP_BIND	1614	1660	NADP.
FT	BINDING	1921	1921	PHOSPHOPANTHETHEINE (BY SIMILARITY).
FT	ACT_SITE	2148	2148	BETA-KETOACYL SYNTHASE.
FT	ACT_SITE	2598	2598	ACYL-ENZYME INTERMEDIATE.
FT	NP_BIND	3058	3104	NADP.
FT	BINDING	3367	3367	PHOSPHOPANTHETHEINE (BY SIMILARITY).
SO	SEQUENCE	3491 AA;	365022 MW;	682BFC32C90F8C4 CRC64;
Query Match	Similarity	40.8%;	Score 173;	DB 1; Length 3491;
Matches	38; Conservative	13;	Mismatches	31; Indels 4; Gaps 1;
Db	1	LTALSP7QOQSLLDLVRAHTMAVINDGNERITASDAPSPASFAHLGFDVYGVYELNRL	60	
Db	1876	LASLPAPREERELPELVSHAAVLGHASABRVAPD-----QAFALGVDSLSALELRRL	1933	
Qy	61	SKATGLRLPVLIPDHTTAAVRL	86	
Db	1932	GAATGVRLPTTVFDPHDPVRTLAHL	1957	
RESULT 3				
OL56_STRAT	STANDARD;	PRT;	3519 AA.	
ID	OL56_STRAT			
AC	Q07017;			
DT	01-NOV-1997 (Rel. 35, Last Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Oleandomycin polyketide synthase, modules 5 and 6.			
GN	ORF6.			
OS	Bacteriomyces antibioticus.			
OC	Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;			
CC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.			
CC	NCB1_taxid=1890;			
CC	[1]			
CC	SEQUENCE FROM N.A.			
CC	MEDLINE=94150470; Pubmed=8107683;			
CC	Swan D.G., Rodriguez A.M., Villache C., Mendez C., Salas J.A.;			
CC	"Characterization of a Streptomycetes antibiotic gene encoding a type			
CC	I polyketide synthase which has an unusual coding sequence."			
CC	Mol. Gen. Genet. 242:358-362(1994).			
CC	-1- FUNCTION: MAY BE INVOLVED IN THE BIOSYNTHESIS OF THE OLEANDOMYCIN			
CC	LACTONE RING.			
CC	-1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETHEINES.			
CC	-1- SIMILARITY: CONTAINS 2 ACTI CARRIER DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.ebi.ac.uk/announcements).			
CC	or send an email to license@ebi.ac.uk .			
CC	-----			
CC	EMBL; L09654; IAA19695.1; -			
DR	HSPB; P25715; IMA.			
DR	InterPro; IPR001227; AC_transferase.			
DR	InterPro; IPR000794; Ketoacyl-synth.			
DR	InterPro; IPR003880; Pantine_acyl.			
DR	InterPro; IPR001031; Thioesterase.			
DR	Pfam; PF00109; ketoacyl-synth. 2.			
DR	Pfam; PF00550; pp-binding; 2.			
DR	Pfam; PF00698; Acyl_transf; 2.			
DR	Pfam; PF00975; Thioesterase; 1.			
DR	Pfam; PF02801; ketoacyl-synth C. 2.			
DR	PROSITE; PS00012; PHOSPHOPANTHETHEINE. 2.			
DR	PROSITE; PS00606; B_KETOACYL_SYNTHASE. 2.			
DR	PROSITE; PS00975; ACYL_DOMAIN; 2.			
KM	Transferase; Acyltransferase; Antibiotic biosynthesis; NADP;			

[illegible]

RT Cloning of the structural gene, sequence analysis and inferred domain
 RL structure of the multifunctional enzyme.";
 CC Eur. J. Biochem. 204:39-49(1992).
 CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
 CC deoxyerythronolide B.
 CC -1- COFACTOR: NADP. CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.
 CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
 CC BIOSYNTHESIS.
 CC -1- MISCELLANEOUS: IN EACH ONE OF ERYA TWO MODULES ARE PRESENT EACH
 CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3
 CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
 CC SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS
 CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
 CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
 CC RESPECTIVELY.
 CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES: ACYLTRANSFERASE (AT),
 CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACTYL CARRIER
 CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
 CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
 CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
 CC OF THE FULL-LENGTH CHAIN.
 CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
 CC -1- SIMILARITY: CONTAINS 2 ACTYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; X56107; CAA39583.1; -
 CC DR EMBL; X63677; AAA26495.1; -
 CC DR EMBL; X63569; CAA44449.1; -
 CC DR HSSP; P00101; ICCH.
 CC DR InterPro; IPR002198; ADH_short.
 CC DR InterPro; IPR001227; Ac transferase.
 CC DR InterPro; IPR000794; Ketoacyl-synt.
 CC DR InterPro; IPR003880; Ppantne attach.
 CC DR InterPro; IPR001031; Thioesterase.
 CC DR Pfam; PF00106; adh_short; 1.
 CC DR Pfam; PF00109; ketoacyl-synt; 2.
 CC DR Pfam; PF00550; pp-binding; 2.
 CC DR Pfam; PF00698; Acyl transf; 2.
 CC DR Pfam; PF00975; Thioesterase; 1.
 CC DR Pfam; PF02801; ketoacyl-synt C; 2.
 CC DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
 CC DR PROSITE; PS5006; B KETOACYL SYNTHASE; 2.
 CC DR TRANSFERASE; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 CC KW Phosphopantetheine; Multifunctional enzyme.
 CC FT DOMAIN 1 1484
 CC FT DOMAIN 1 1485 3172
 CC FT DOMAIN 1 1486 3172
 CC FT DOMAIN 1 1487 3172
 CC FT DOMAIN 1 1488 3172
 CC FT DOMAIN 1 1489 3172
 CC FT DOMAIN 1 1490 3172
 CC FT DOMAIN 1 1491 3172
 CC FT DOMAIN 1 1492 3172
 CC FT DOMAIN 1 1493 3172
 CC FT DOMAIN 1 1494 3172
 CC FT DOMAIN 1 1495 3172
 CC FT DOMAIN 1 1496 3172
 CC FT DOMAIN 1 1497 3172
 CC FT DOMAIN 1 1498 3172
 CC FT DOMAIN 1 1499 3172
 CC FT DOMAIN 1 1500 3172
 CC FT DOMAIN 1 1501 3172
 CC FT DOMAIN 1 1502 3172
 CC FT DOMAIN 1 1503 3172
 CC FT DOMAIN 1 1504 3172
 CC FT DOMAIN 1 1505 3172
 CC FT DOMAIN 1 1506 3172
 CC FT DOMAIN 1 1507 3172
 CC FT DOMAIN 1 1508 3172
 CC FT DOMAIN 1 1509 3172
 CC FT DOMAIN 1 1510 3172
 CC FT DOMAIN 1 1511 3172
 CC FT DOMAIN 1 1512 3172
 CC FT DOMAIN 1 1513 3172
 CC FT DOMAIN 1 1514 3172
 CC FT DOMAIN 1 1515 3172
 CC FT DOMAIN 1 1516 3172
 CC FT DOMAIN 1 1517 3172
 CC FT DOMAIN 1 1518 3172
 CC FT DOMAIN 1 1519 3172
 CC FT DOMAIN 1 1520 3172
 CC FT DOMAIN 1 1521 3172
 CC FT DOMAIN 1 1522 3172
 CC FT DOMAIN 1 1523 3172
 CC FT DOMAIN 1 1524 3172
 CC FT DOMAIN 1 1525 3172
 CC FT DOMAIN 1 1526 3172
 CC FT DOMAIN 1 1527 3172
 CC FT DOMAIN 1 1528 3172
 CC FT DOMAIN 1 1529 3172
 CC FT DOMAIN 1 1530 3172
 CC FT DOMAIN 1 1531 3172
 CC FT DOMAIN 1 1532 3172
 CC FT DOMAIN 1 1533 3172
 CC FT DOMAIN 1 1534 3172
 CC FT DOMAIN 1 1535 3172
 CC FT DOMAIN 1 1536 3172
 CC FT DOMAIN 1 1537 3172
 CC FT DOMAIN 1 1538 3172
 CC FT DOMAIN 1 1539 3172
 CC FT DOMAIN 1 1540 3172
 CC FT DOMAIN 1 1541 3172
 CC FT DOMAIN 1 1542 3172
 CC FT DOMAIN 1 1543 3172
 CC FT DOMAIN 1 1544 3172
 CC FT DOMAIN 1 1545 3172
 CC FT DOMAIN 1 1546 3172
 CC FT DOMAIN 1 1547 3172
 CC FT DOMAIN 1 1548 3172
 CC FT DOMAIN 1 1549 3172
 CC FT DOMAIN 1 1550 3172
 CC FT DOMAIN 1 1551 3172
 CC FT DOMAIN 1 1552 3172
 CC FT DOMAIN 1 1553 3172
 CC FT DOMAIN 1 1554 3172
 CC FT DOMAIN 1 1555 3172
 CC FT DOMAIN 1 1556 3172
 CC FT DOMAIN 1 1557 3172
 CC FT DOMAIN 1 1558 3172
 CC FT DOMAIN 1 1559 3172
 CC FT DOMAIN 1 1560 3172
 CC FT DOMAIN 1 1561 3172
 CC FT DOMAIN 1 1562 3172
 CC FT DOMAIN 1 1563 3172
 CC FT DOMAIN 1 1564 3172
 CC FT DOMAIN 1 1565 3172
 CC FT DOMAIN 1 1566 3172
 CC FT DOMAIN 1 1567 3172
 CC FT DOMAIN 1 1568 3172
 CC FT DOMAIN 1 1569 3172
 CC FT DOMAIN 1 1570 3172
 CC FT DOMAIN 1 1571 3172
 CC FT DOMAIN 1 1572 3172
 CC FT DOMAIN 1 1573 3172
 CC FT DOMAIN 1 1574 3172
 CC FT DOMAIN 1 1575 3172
 CC FT DOMAIN 1 1576 3172
 CC FT DOMAIN 1 1577 3172
 CC FT DOMAIN 1 1578 3172
 CC FT DOMAIN 1 1579 3172
 CC FT DOMAIN 1 1580 3172
 CC FT DOMAIN 1 1581 3172
 CC FT DOMAIN 1 1582 3172
 CC FT DOMAIN 1 1583 3172
 CC FT DOMAIN 1 1584 3172
 CC FT DOMAIN 1 1585 3172
 CC FT DOMAIN 1 1586 3172
 CC FT DOMAIN 1 1587 3172
 CC FT DOMAIN 1 1588 3172
 CC FT DOMAIN 1 1589 3172
 CC FT DOMAIN 1 1590 3172
 CC FT DOMAIN 1 1591 3172
 CC FT DOMAIN 1 1592 3172
 CC FT DOMAIN 1 1593 3172
 CC FT DOMAIN 1 1594 3172
 CC FT DOMAIN 1 1595 3172
 CC FT DOMAIN 1 1596 3172
 CC FT DOMAIN 1 1597 3172
 CC FT DOMAIN 1 1598 3172
 CC FT DOMAIN 1 1599 3172
 CC FT DOMAIN 1 1600 3172
 CC FT DOMAIN 1 1601 3172
 CC FT DOMAIN 1 1602 3172
 CC FT DOMAIN 1 1603 3172
 CC FT DOMAIN 1 1604 3172
 CC FT DOMAIN 1 1605 3172
 CC FT DOMAIN 1 1606 3172
 CC FT DOMAIN 1 1607 3172
 CC FT DOMAIN 1 1608 3172
 CC FT DOMAIN 1 1609 3172
 CC FT DOMAIN 1 1610 3172
 CC FT DOMAIN 1 1611 3172
 CC FT DOMAIN 1 1612 3172
 CC FT DOMAIN 1 1613 3172
 CC FT DOMAIN 1 1614 3172
 CC FT DOMAIN 1 1615 3172
 CC FT DOMAIN 1 1616 3172
 CC FT DOMAIN 1 1617 3172
 CC FT DOMAIN 1 1618 3172
 CC FT DOMAIN 1 1619 3172
 CC FT DOMAIN 1 1620 3172
 CC FT DOMAIN 1 1621 3172
 CC FT DOMAIN 1 1622 3172
 CC FT DOMAIN 1 1623 3172
 CC FT DOMAIN 1 1624 3172
 CC FT DOMAIN 1 1625 3172
 CC FT DOMAIN 1 1626 3172
 CC FT DOMAIN 1 1627 3172
 CC FT DOMAIN 1 1628 3172
 CC FT DOMAIN 1 1629 3172
 CC FT DOMAIN 1 1630 3172
 CC FT DOMAIN 1 1631 3172
 CC FT DOMAIN 1 1632 3172
 CC FT DOMAIN 1 1633 3172
 CC FT DOMAIN 1 1634 3172
 CC FT DOMAIN 1 1635 3172
 CC FT DOMAIN 1 1636 3172
 CC FT DOMAIN 1 1637 3172
 CC FT DOMAIN 1 1638 3172
 CC FT DOMAIN 1 1639 3172
 CC FT DOMAIN 1 1640 3172
 CC FT DOMAIN 1 1641 3172
 CC FT DOMAIN 1 1642 3172
 CC FT DOMAIN 1 1643 3172
 CC FT DOMAIN 1 1644 3172
 CC FT DOMAIN 1 1645 3172
 CC FT DOMAIN 1 1646 3172
 CC FT DOMAIN 1 1647 3172
 CC FT DOMAIN 1 1648 3172
 CC FT DOMAIN 1 1649 3172
 CC FT DOMAIN 1 1650 3172
 CC FT DOMAIN 1 1651 3172
 CC FT DOMAIN 1 1652 3172
 CC FT DOMAIN 1 1653 3172
 CC FT DOMAIN 1 1654 3172
 CC FT DOMAIN 1 1655 3172
 CC FT DOMAIN 1 1656 3172
 CC FT DOMAIN 1 1657 3172
 CC FT DOMAIN 1 1658 3172
 CC FT DOMAIN 1 1659 3172
 CC FT DOMAIN 1 1660 3172
 CC FT DOMAIN 1 1661 3172
 CC FT DOMAIN 1 1662 3172
 CC FT DOMAIN 1 1663 3172
 CC FT DOMAIN 1 1664 3172
 CC FT DOMAIN 1 1665 3172
 CC FT DOMAIN 1 1666 3172
 CC FT DOMAIN 1 1667 3172
 CC FT DOMAIN 1 1668 3172
 CC FT DOMAIN 1 1669 3172
 CC FT DOMAIN 1 1670 3172
 CC FT DOMAIN 1 1671 3172
 CC FT DOMAIN 1 1672 3172
 CC FT DOMAIN 1 1673 3172
 CC FT DOMAIN 1 1674 3172
 CC FT DOMAIN 1 1675 3172
 CC FT DOMAIN 1 1676 3172
 CC FT DOMAIN 1 1677 3172
 CC FT DOMAIN 1 1678 3172
 CC FT DOMAIN 1 1679 3172
 CC FT DOMAIN 1 1680 3172
 CC FT DOMAIN 1 1681 3172
 CC FT DOMAIN 1 1682 3172
 CC FT DOMAIN 1 1683 3172
 CC FT DOMAIN 1 1684 3172
 CC FT DOMAIN 1 1685 3172
 CC FT DOMAIN 1 1686 3172
 CC FT DOMAIN 1 1687 3172
 CC FT DOMAIN 1 1688 3172
 CC FT DOMAIN 1 1689 3172
 CC FT DOMAIN 1 1690 3172
 CC FT DOMAIN 1 1691 3172
 CC FT DOMAIN 1 1692 3172
 CC FT DOMAIN 1 1693 3172
 CC FT DOMAIN 1 1694 3172
 CC FT DOMAIN 1 1695 3172
 CC FT DOMAIN 1 1696 3172
 CC FT DOMAIN 1 1697 3172
 CC FT DOMAIN 1 1698 3172
 CC FT DOMAIN 1 1699 3172
 CC FT DOMAIN 1 1700 3172
 CC FT DOMAIN 1 1701 3172
 CC FT DOMAIN 1 1702 3172
 CC FT DOMAIN 1 1703 3172
 CC FT DOMAIN 1 1704 3172
 CC FT DOMAIN 1 1705 3172
 CC FT DOMAIN 1 1706 3172
 CC FT DOMAIN 1 1707 3172
 CC FT DOMAIN 1 1708 3172
 CC FT DOMAIN 1 1709 3172
 CC FT DOMAIN 1 1710 3172
 CC FT DOMAIN 1 1711 3172
 CC FT DOMAIN 1 1712 3172
 CC FT DOMAIN 1 1713 3172
 CC FT DOMAIN 1 1714 3172
 CC FT DOMAIN 1 1715 3172
 CC FT DOMAIN 1 1716 3172
 CC FT DOMAIN 1 1717 3172
 CC FT DOMAIN 1 1718 3172
 CC FT DOMAIN 1 1719 3172
 CC FT DOMAIN 1 1720 3172
 CC FT DOMAIN 1 1721 3172
 CC FT DOMAIN 1 1722 3172
 CC FT DOMAIN 1 1723 3172
 CC FT DOMAIN 1 1724 3172
 CC FT DOMAIN 1 1725 3172
 CC FT DOMAIN 1 1726 3172
 CC FT DOMAIN 1 1727 3172
 CC FT DOMAIN 1 1728 3172
 CC FT DOMAIN 1 1729 3172
 CC FT DOMAIN 1 1730 3172
 CC FT DOMAIN 1 1731 3172
 CC FT DOMAIN 1 1732 3172
 CC FT DOMAIN 1 1733 3172
 CC FT DOMAIN 1 1734 3172
 CC FT DOMAIN 1 1735 3172
 CC FT DOMAIN 1 1736 3172
 CC FT DOMAIN 1 1737 3172
 CC FT DOMAIN 1 1738 3172
 CC FT DOMAIN 1 1739 3172
 CC FT DOMAIN 1 1740 3172
 CC FT DOMAIN 1 1741 3172
 CC FT DOMAIN 1 1742 3172
 CC FT DOMAIN 1 1743 3172
 CC FT DOMAIN 1 1744 3172
 CC FT DOMAIN 1 1745 3172
 CC FT DOMAIN 1 1746 3172
 CC FT DOMAIN 1 1747 3172
 CC FT DOMAIN 1 1748 3172
 CC FT DOMAIN 1 1749 3172
 CC FT DOMAIN 1 1750 3172
 CC FT DOMAIN 1 1751 3172
 CC FT DOMAIN 1 1752 3172
 CC FT DOMAIN 1 1753 3172
 CC FT DOMAIN 1 1754 3172
 CC FT DOMAIN 1 1755 3172
 CC FT DOMAIN 1 1756 3172
 CC FT DOMAIN 1 1757 3172
 CC FT DOMAIN 1 1758 3172
 CC FT DOMAIN 1 1759 3172
 CC FT DOMAIN 1 1760 3172
 CC FT DOMAIN 1 1761 3172
 CC FT DOMAIN 1 1762 3172
 CC FT DOMAIN 1 1763 3172
 CC FT DOMAIN 1 1764 3172
 CC FT DOMAIN 1 1765 3172
 CC FT DOMAIN 1 1766 3172
 CC FT DOMAIN 1 1767 3172
 CC FT DOMAIN 1 1768 3172
 CC FT DOMAIN 1 1769 3172
 CC FT DOMAIN 1 1770 3172
 CC FT DOMAIN 1 1771 3172
 CC FT DOMAIN 1 1772 3172
 CC FT DOMAIN 1 1773 3172
 CC FT DOMAIN 1 1774 3172
 CC FT DOMAIN 1 1775 3172
 CC FT DOMAIN 1 1776 3172
 CC FT DOMAIN 1 1777 3172
 CC FT DOMAIN 1 1778 3172
 CC FT DOMAIN 1 1779 3172
 CC FT DOMAIN 1 1780 3172
 CC FT DOMAIN 1 1781 3172
 CC FT DOMAIN 1 1782 3172
 CC FT DOMAIN 1 1783 3172
 CC FT DOMAIN 1 1784 3172
 CC FT DOMAIN 1 1785 3172
 CC FT DOMAIN 1 1786 3172
 CC FT DOMAIN 1 1787 3172
 CC FT DOMAIN 1 1788 3172
 CC FT DOMAIN 1 1789 3172
 CC FT DOMAIN 1 1790 3172
 CC FT DOMAIN 1 1791 3172
 CC FT DOMAIN 1 1792 3172
 CC FT DOMAIN 1 1793 3172
 CC FT DOMAIN 1 1794 3172
 CC FT DOMAIN 1 1795 3172
 CC FT DOMAIN 1 1796 3172
 CC FT DOMAIN 1 1797 3172
 CC FT DOMAIN 1 1798 3172
 CC FT DOMAIN 1 1799 3172
 CC FT DOMAIN 1 1800 3172
 CC FT DOMAIN 1 1801 3172
 CC FT DOMAIN 1 1802 3172
 CC FT DOMAIN 1 1803 3172
 CC FT DOMAIN 1 1804 3172
 CC FT DOMAIN 1 1805 3172
 CC FT DOMAIN 1 1806 3172
 CC FT DOMAIN 1 1807 3172
 CC FT DOMAIN 1 1808 3172
 CC FT DOMAIN 1 1809 3172
 CC FT DOMAIN 1 1810 3172
 CC FT DOMAIN 1 1811 3172
 CC FT DOMAIN 1 1812 3172
 CC FT DOMAIN 1 1813 3172
 CC FT DOMAIN 1 1814 3172
 CC FT DOMAIN 1 1815 3172
 CC FT DOMAIN 1 1816 3172
 CC FT DOMAIN 1 1817 3172
 CC FT DOMAIN 1 1818 3172
 CC FT DOMAIN 1 1819 3172
 CC FT DOMAIN 1 1820 3172
 CC FT DOMAIN 1 1821 3172
 CC FT DOMAIN 1 1822 3172
 CC FT DOMAIN 1 1823 3172
 CC FT DOMAIN 1 1824 3172
 CC FT DOMAIN 1 1825 3172
 CC FT DOMAIN 1 1826 3172
 CC FT DOMAIN 1 1827 3172
 CC FT DOMAIN 1 1828 3172
 CC FT DOMAIN 1 1829 3172
 CC FT DOMAIN 1 1830 3172
 CC FT DOMAIN 1 1831 3172
 CC FT DOMAIN 1 1832 3172
 CC FT DOMAIN 1 1833 3172
 CC FT DOMAIN 1 1834 3172
 CC FT DOMAIN 1 1835 3172
 CC FT DOMAIN 1 1836 3172
 CC FT DOMAIN 1 1837 3172
 CC FT DOMAIN 1 1838 3172
 CC FT DOMAIN 1 1839 3172
 CC FT DOMAIN 1 1840 3172
 CC FT DOMAIN 1 1841 3172
 CC FT DOMAIN 1 1842 3172
 CC FT DOMAIN 1 1843 3172
 CC FT DOMAIN 1 1844 3172
 CC FT DOMAIN 1 1845 3172
 CC FT DOMAIN 1 1846 3172
 CC FT DOMAIN 1 1847 3172
 CC FT DOMAIN 1 1848 3172
 CC FT DOMAIN 1 1849 3172
 CC FT DOMAIN 1 1850 3172
 CC FT DOMAIN 1 1851 3172
 CC FT DOMAIN 1 1852 3172
 CC FT DOMAIN 1 1853 3172
 CC FT DOMAIN 1 1854 3172
 CC FT DOMAIN 1 1855 3172
 CC FT DOMAIN 1 1856 3172
 CC FT DOMAIN 1 1857 3172
 CC FT DOMAIN 1 1858 3172
 CC FT DOMAIN 1 1859 3172
 CC FT DOMAIN 1 1860 3172
 CC FT DOMAIN 1 1861 3172
 CC FT DOMAIN 1 1862 3172
 CC FT DOMAIN 1 1863 3172
 CC FT DOMAIN 1 1864 3172
 CC FT DOMAIN 1 1865 3172
 CC FT DOMAIN 1 1866 3172
 CC FT DOMAIN 1 1867 3172
 CC FT DOMAIN 1 1868 3172
 CC FT DOMAIN 1 1869 3172
 CC FT DOMAIN 1 1870 3172
 CC FT DOMAIN 1 1871 3172
 CC FT DOMAIN 1 1872 3172
 CC FT DOMAIN 1 1873 3172
 CC FT DOMAIN 1 1874 3172
 CC FT DOMAIN 1 1875 3172
 CC FT DOMAIN 1 1876 3172
 CC FT DOMAIN 1 1877 3172
 CC FT DOMAIN 1 1878 3172
 CC FT DOMAIN 1 1879 3172
 CC FT DOMAIN 1 1880 3172
 CC FT DOMAIN 1 1881 3172
 CC FT DOMAIN 1 1882 3172
 CC FT DOMAIN 1 1883 3172
 CC FT DOMAIN 1 1884 3172
 CC FT DOMAIN 1 1885 3172
 CC FT DOMAIN 1 1886 3172
 CC FT DOMAIN 1 1887 3172
 CC FT DOMAIN 1 1888 3172
 CC FT DOMAIN 1 1889 3172
 CC FT DOMAIN 1 1890 3172
 CC FT DOMAIN 1 1891 3172
 CC FT DOMAIN 1 1892 3172
 CC FT DOMAIN 1 1893 3172
 CC FT DOMAIN 1 1894 3172
 CC FT DOMAIN 1 1895 3172
 CC FT DOMAIN 1 1896 3172
 CC FT DOMAIN 1 1897 3172
 CC FT DOMAIN 1 1898 3172
 CC FT DOMAIN 1 1899 3172
 CC FT DOMAIN 1 1900 3172
 CC FT DOMAIN 1 1901 3172
 CC FT DOMAIN 1 1902 3172
 CC FT DOMAIN 1 1903 3172
 CC FT DOMAIN 1 1904 3172
 CC FT DOMAIN 1 1905 3172
 CC FT DOMAIN 1 1906 3172
 CC FT DOMAIN 1 1907 3172
 CC FT DOMAIN 1 1908 3172
 CC FT DOMAIN 1 1909 3172
 CC FT DOMAIN 1 1910 3172
 CC FT DOMAIN 1 1911 3172
 CC FT DOMAIN 1 1912 3172
 CC FT DOMAIN 1 1913 3172
 CC FT DOMAIN 1 1914 3172
 CC FT DOMAIN 1 1915 3172
 CC FT DOMAIN 1 1916 3172
 CC FT DOMAIN 1 1917 3172
 CC FT DOMAIN 1 1918 3172
 CC FT DOMAIN 1 1919 3172
 CC FT DOMAIN 1 1920 3172
 CC FT DOMAIN 1 1921 3172
 CC FT DOMAIN 1 1922 3172
 CC FT DOMAIN 1 1923 3172
 CC FT DOMAIN 1 1924 3172
 CC FT DOMAIN 1 1925 3172
 CC FT DOMAIN 1 1926 3172
 CC FT DOMAIN 1 1927 3172
 CC FT DOMAIN 1 1928 3172
 CC FT DOMAIN 1 1929 3172
 CC FT DOMAIN 1 1930 3172
 CC FT DOMAIN 1 1931 3172
 CC FT DOMAIN 1 1932 3172
 CC FT DOMAIN 1 1933 3172
 CC FT DOMAIN 1 1934 3172
 CC FT DOMAIN 1 1935 3172
 CC FT DOMAIN 1 1936 3172
 CC FT DOMAIN 1 1937 3172
 CC FT DOMAIN 1 1938 3172
 CC FT DOMAIN 1 1939 3172
 CC FT DOMAIN 1 1940 3172
 CC FT DOMAIN 1 1941 3172
 CC FT DOMAIN 1 1942 3172
 CC FT DOMAIN 1 1943 3172
 CC FT DOMAIN 1 1944 3172
 CC FT DOMAIN 1 1945 3172
 CC FT DOMAIN 1 1946 3172
 CC FT DOMAIN 1 1947 3172
 CC FT DOMAIN 1 1948 3172
 CC FT DOMAIN 1 1949 3172
 CC FT DOMAIN 1 1950 3172
 CC FT DOMAIN 1 1951 3172
 CC FT DOMAIN 1 1952 3172
 CC FT DOMAIN 1 1953 3172
 CC FT DOMAIN 1 1954 3172
 CC FT DOMAIN 1 1955 3172
 CC FT DOMAIN 1 1956 3172
 CC FT DOMAIN 1 1957 3172
 CC FT DOMAIN 1 1958 3172
 CC FT DOMAIN 1 1959 3172
 CC FT DOMAIN 1 1960 3172
 CC FT DOMAIN 1 1961 3172
 CC FT DOMAIN 1 1962 3172
 CC FT DOMAIN 1 1963 3172
 CC FT DOMAIN 1 1964 3172
 CC FT DOMAIN 1 1965 3172
 CC FT DOMAIN 1 1966 3172
 CC FT DOMAIN 1 1967 3172
 CC FT DOMAIN 1 1968 3172
 CC FT DOMAIN 1 1969 3172
 CC FT DOMAIN 1 1970 3172
 CC FT DOMAIN 1 1971 3172
 CC FT DOMAIN 1 1972 3172
 CC FT DOMAIN 1 1973 3172
 CC FT DOMAIN 1 1974 3172
 CC FT DOMAIN 1 1975 3172
 CC FT DOMAIN 1 1976 3172
 CC FT DOMAIN 1 1977 3172
 CC FT DOMAIN 1 1978 3172
 CC FT DOMAIN 1 1979 3172
 CC FT DOMAIN 1 1980 3172
 CC FT DOMAIN 1 1981 3172
 CC FT DOMAIN 1 1982 3172
 CC FT DOMAIN 1 1983 3172
 CC FT DOMAIN 1 1984 3172
 CC FT DOMAIN 1 1985 3172
 CC FT DOMAIN 1 1986 3172
 CC FT DOMAIN 1 1987 3172
 CC FT DOMAIN 1 1988 3172
 CC FT DOMAIN 1 1989 3172
 CC FT DOMAIN 1 1990 3172
 CC FT DOMAIN 1 1991 3172
 CC FT DOMAIN 1 1992 3172
 CC FT DOMAIN 1 1993 3172
 CC FT DOMAIN 1 1994 3172
 CC FT DOMAIN 1 1995 3172
 CC FT DOMAIN 1 1996 3172
 CC FT DOMAIN 1 1997 3172
 CC FT DOMAIN 1 1998 3172
 CC FT DOMAIN 1 1999 3172
 CC FT DOMAIN 1 2000 3172
 CC FT DOMAIN 1 2001 3172
 CC FT DOMAIN 1 2002 3172
 CC FT DOMAIN 1 2003 3172
 CC FT DOMAIN 1 2004 3172
 CC FT DOMAIN 1 2005 3172
 CC FT DOMAIN 1 2006 3172
 CC FT DOMAIN 1 2007 3172
 CC FT DOMAIN 1 2008 3172
 CC FT DOMAIN 1 2009 3172
 CC FT DOMAIN 1 2010 3172
 CC FT DOMAIN 1 2011 3172
 CC FT DOMAIN 1 2012 3172
 CC FT DOMAIN 1 2013 3172
 CC FT DOMAIN 1 2014 3172
 CC FT DOMAIN 1 2015 3172
 CC FT DOMAIN 1 2016 3172
 CC FT DOMAIN 1 2017 3172
 CC FT DOMAIN 1 2018 3172
 CC FT DOMAIN 1 2019 3172
 CC FT DOMAIN 1 2020 3172
 CC FT DOMAIN 1 2021 3172
 CC FT DOMAIN 1 2022 3172
 CC FT DOMAIN 1 2023 3172
 CC FT DOMAIN 1 2024 3172
 CC FT DOMAIN 1 2025 3172
 CC FT DOMAIN 1 2026 3172
 CC FT DOMAIN 1 2027 3172
 CC FT DOMAIN 1 2028 3172
 CC FT DOMAIN 1 2029 3172
 CC FT DOMAIN 1 2030 3172
 CC FT DOMAIN 1 2031 3172
 CC FT DOMAIN 1 2032 3172
 CC FT DOMAIN 1 2033 3172
 CC FT DOMAIN 1 2034 3172
 CC FT DOMAIN 1 2035 3172
 CC FT DOMAIN 1 2036 3172
 CC FT DOMAIN 1 2037 3172
 CC FT DOMAIN 1 2038 3172
 CC FT DOMAIN 1 2039 3172
 CC FT DOMAIN 1 2040 3172
 CC FT DOMAIN 1 2041 3172
 CC FT DOMAIN 1 2042 3172
 CC FT DOMAIN 1 2043 3172
 CC FT DOMAIN 1 2044 3172
 CC FT DOMAIN 1 2045 3172
 CC FT DOMAIN 1 2046 3172
 CC FT DOMAIN 1 2047 3172
 CC FT DOMAIN 1 2048 3172
 CC FT DOMAIN 1 2049 3172
 CC FT DOMAIN 1 2050 3172
 CC FT DOMAIN 1 2051 3172
 CC FT DOMAIN 1 2052 3172
 CC FT DOMAIN 1 2053 3172
 CC FT DOMAIN 1 2054 3172
 CC FT DOMAIN 1 2055 3172
 CC FT DOMAIN 1 2056 3172
 CC FT DOMAIN 1 2057 3172
 CC FT DOMAIN 1 2058 3172
 CC FT DOMAIN 1 2059 3172
 CC FT DOMAIN 1 2060 3172
 CC FT DOMAIN 1 2061 3172
 CC FT DOMAIN 1 2062 317

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentile S., Hamlin N., Holroyd S.,
 RA Hornby T., Jaseil K., Ketchum A., McLean J., Moul S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Ratandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Suleman J.E., Taylor K., Whitehead S., Barrell B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Blahut W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLPTIICEROL
 CC SYNTHESIS.
 CC -1- CORPACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETHEINE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AL021070; CA15929.1; -
 DR EMBL: AE007122; AA47329.1; -
 DR TIGR: MT3002; -
 DR TubercuList; RV2932; -
 DR InterPro: IPR001227; AC transferase.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; pantoic attach.
 DR Pfam: PF00109; ketoacyl-synt. 1.
 DR Pfam: PF00550; pp-binding; 1.
 DR Pfam: PF00698; Acyl transfer; 1.
 DR Pfam: PF02801; ketoacyl-synt C; 1.
 DR PROSITE: PS00606; B KETOACYL SYNTHASE; 1.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; FALSE_NEG.
 DR PROSITE: PS50075; ACP DOMAIN; 1.
 DR Hypothetical protein; Multifunctional enzyme; Oxidoreductase;
 KM transferase; NADP; Phosphopantetheine; Complete proteome.
 FT DOMAIN 1425 1495 ACYL CARRIER (ACP).
 FT BINDING 1458 1458 PHOSPHOPANTHETHEINE (POTENTIAL).
 FT SEQUENCE 1538 AA; 162527 MW; B55E2A2042AD00CC CRC64;
 SQ
 Query Match 23.3%; Score 99; DB 1; Length 1538;
 Best Local Similarity 35.2%; Pred. No. 0.0029;
 Matches 25; Conservative 13; Mismatches 29; Indels 4; Gaps 1;
 QY 8 QOQSLLDLVRAHTAVLVNDGNERTRASDAGSPASFAHGFSDVMGVELRNRLSKATGLR 67
 DB 1420 KRDDMFDFHVGALAAATVGMPTPE---PLDPSAGFGFOLQMDLSIMVTLORALSISIGER 1475
 QY 68 LPTVLIFFDHTT 78
 DB 1476 LPASVVPDYPT 1486
 MSAS_PENPA STANDARD; PRT; 1774 AA.
 ID P22367;
 AC 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 15-UTN-2002 (Rel. 41, Last annotation update)
 DE 6-methylsalicylic acid synthase (EC 2.3.1.-) (MSAS).
 OS Penticillium patulum (Penticillium griseofulvum).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; microsporitic Trichocomaceae; Penticillium.
 OK NCBI_Taxid=5078;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=DSM 62662;
 RX MEDLINE=91006137; PubMed=2209605;
 RA Beck U., Ripka S., Stegner A., Schiltz E., Schweizer E.;
 RT "The multifunctional 6-methylsalicylic acid synthase gene of
 RT Penticillium patulum. Its gene structure relative to that of other
 RT polyketide synthases.";
 RL Eur. J. Biochem. 192:487-498(1990).
 CC -1- FUNCTION: This multifunctional enzyme is a polyketide synthase.
 CC It catalyzes a total of 11 steps by seven different component
 CC enzymes, in the biosynthesis of the antibiotic patulin.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + 3 malonyl-CoA + NADPH =
 CC 6-methylsalicylic acid + NADP(+) + 3 CO(2) + H(2)O.
 CC -1- PATHWAY: Patulin biosynthesis.
 CC -1- SUBUNIT: HOMOMULTIMER.
 CC -1- INDUCTION: IN THE LATE LOGARITHMIC GROWTH PHASE.
 CC -1- SIMILARITY: WITH MOST POLYKETIDE SYNTHASES, SPECIALLY WITH RAT
 CC FATTY ACID SYNTHASE, AND WITH OTHER ENZYMES SUCH AS LIPASES AND
 CC THIOLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X55776; CA19295.1; -
 DR PIR: S13178; S13178.
 DR InterPro: IPR001227; AC transferase.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; pantoic attach.
 DR Pfam: PF00109; ketoacyl-synt. 1.
 DR Pfam: PF00550; pp-binding; 1.
 DR Pfam: PF00698; Acyl transfer; 1.
 DR Pfam: PF02801; ketoacyl-synt C; 1.
 DR PROSITE: PS00606; B KETOACYL SYNTHASE; 1.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 1.
 DR PROSITE: PS50075; ACP DOMAIN; 1.
 DR Multifunctional enzyme; Antibiotic biosynthesis;
 KM transferase; NADP; Phosphopantetheine;
 FT DOMAIN 166 238 ACYLTRANSFERASE (AT).
 FT DOMAIN 642 676 ACETYL/MALONYL TRANSFERASES.
 FT DOMAIN 1403 1450 2-OXOACYL REDUCTASE.
 FT DOMAIN 1700 1769 ACYL CARRIER (ACP).
 FT NE BIND 1419 1424 NADP (POTENTIAL).
 FT ACT SITE 204 204 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT SITE 653 653 MALONYLTRANSFERASE (BY SIMILARITY).
 FT BINDING 1732 1732 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT SEQUENCE 1774 AA; 190732 MW; 05ED5DD10863F938 CRC64;
 SQ
 Query Match 21.3%; Score 90.5; DB 1; Length 1774;
 Best Local Similarity 32.9%; Pred. No. 0.033;
 Matches 26; Conservative 13; Mismatches 35; Indels 5; Gaps 1;
 QY 5 SPTQOOSLLDLVRAHTAVLVNDGNERTRASDAGSPASFAHGFSDVMGVELRNRLSKATGLR 64
 DB 1692 SGBELKAYLDKIRGCVARVL-----QMTABDVDSKALADLVGVSVMVTLRRQLQTLR 1746
 QY 65 GLRLPVTLLIFDHTT 83
 DB 1747 KIAVPELTWMSHTVSHLA 1765
 MSAS_PENPA STANDARD; PRT; 1774 AA.
 ID P22367;
 AC 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)

PKSL_BACSU STANDARD; PRT; 4427 AA.

ID PKSL_BACSU 005470;

AC 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Putative polyketide synthase pksL (PKS).

GN PKSL OR PKSX OR PKSA OR OUTG.

OS Bacillus subtilis.

OC Bacillaria Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168 / PBI424;

RX MEDLINE=93345824; PubMed=8344529;

RA Scotti C., Piatti M., Cuzzoni A., Perani P., Tognoni A., Grandi G.,

RA Galizzi A., Albertini A.M.,

RT "A Bacillus subtilis large ORF coding for a polypeptide highly

RT similar to polyketide synthases."

RL Gene 130:65-71(1993).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogassawa N., Moser I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertorello M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borris R., Bourcier L., Brans A., Braun M., Bridgell S.C., Bron S.,

RA Brouillet S., Brunehi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

RA Denton F., Devine K.M., Duesternof A., Ehrlich S.D., Emmerson P.T.,

RA Entlan K.D., Errington J., Fabre C., Ferrati E., Foulger D.,

RA Fritz C., Fujita Y., Fujita Y., Funa S., Galizzi A., Galleron N.,

RA Gim S.Y., Glaeser P., Goffeau A., Goldlight J.E., Grandi G.,

RA Giuseppe G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holzapfel S., Hosono S., Hulio M.F., Itaya M., Jones L.,

RA Jorib B., Karamata D., Kasahara Y., Kaeer-Blanchard M., Klein C.,

RA Kobayashi Y., Koelter P., Koningstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Paro V., Pohl T.M., Portetelle D., Potworlik S., Prescott A.M.,

RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

RA Sekiuchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Solio B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambuit R., Wedler E., Wedler H., Weitzenecker T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zanchin A.,

RT "The complete genome sequence of the Gram-positive bacterium Bacillus

RT subtilis."

RT Nature 390:249-256(1997).

RL [3]

RP SEQUENCE OF 3619-4427 FROM N.A.

RC STRAIN=168 / PBI424;

RA Grandi G.,

RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR

CC THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN

CC SECONDARY METABOLISM.

CC -1- COFACTOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES

CC (POTENTIAL).

CC -1- SIMILARITY: CONTAINS 5 ACYL CARRIER DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.

CC -----

CC EMBL: 214098; CAAT8479.1; -

DR EMBL: U11039; AA85145.1; -

DR EMBL: 229113; CAB13602.1; -

DR EMBL: 235133; CAB4504.1; -

DR PIR: S25021; S25021.

DR Subtilisin; BG10698; PKSL.

DR InterPro: IPR002198; ADH short.

DR InterPro: IPR000794; Ketoacyl-synt.

DR InterPro: IPR003880; Ppantn2_attach.

DR Pfam: PF00106; adh_short; 1.

DR Pfam: PF00109; ketoacyl-synt; 4.

DR Pfam: PF00550; pp-binding; 5.

DR Pfam: PF02801; ketoacyl-synt C; 4.

DR PROSITE: PS00606; B_KETOACYL-SYNTHASE; 5.

DR PROSITE: PS50075; ACP DOMAIN; 5.

DR TRANSFERASE; Acyltransferase; Antibiotic biosynthesis; NADP;

KW Phosphopantetheine; Multifunctional enzyme; Repeat; Complete proteome.

FT DOMAIN 211 280

FT DOMAIN 382 759

FT DOMAIN 937 1115

FT DOMAIN 1409 1602

FT DOMAIN 1687 1759

FT DOMAIN 1876 2253

FT DOMAIN 2491 2560

FT DOMAIN 2632 2701

FT DOMAIN 2823 3182

FT DOMAIN 3575 3776

FT DOMAIN 3854 3923

FT DOMAIN 4019 4373

FT BINDING 243 243

FT BINDING 1723 1723

FT BINDING 2523 2523

FT BINDING 2664 2664

FT BINDING 3886 3886

SQ SEQUENCE 4427 AA; 493398 MW; 96125218561AB9F2 CRC64;

Query Match 19.9%; Score 84.5; DB 1; Length 4427;

Best local similarity 30.2%; Pred. No. 0.49; Indels 9; Gaps 2;

Matches 26; Conservative 14; Mismatches 37;

QY 5 SPTQOQSL-----LLDLVRAHTAAVINDDGNERTASDAGPSAFHGFDSVWGVEIRNL 60

DB 2479 SPTRETTIAEALCDELTAIGLAFLVMDNE-----IDDEAFIDIGMDSITGLERIKAI 2533

QY 61 SKATGRLPVTLLIFDHTTPAAVAARL 86

DB 2534 NKQYGTSLNVTQVYDPTTRDEAVVL 2559

RESULT 8

PPSA MYCTU

ID PPSA MYCTU STANDARD; PRT; 1876 AA.

AC 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Phenoliphiocetol synthesis polyketide synthase ppsa.

GN PPSA OR RV2931 OR M13000 OR M13000 OR M13000 OR M13000 OR M13000

OS Mycobacterium tuberculosis.

OC Bacillaria; Actinobacteria; Actinobacteria (class); Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;

RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Bresch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmeier K., Gao S., Barry C.E. III, Tekala F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,

RA Hornaby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger J., Skellton S., Squares S., Squares R.,
 RA Sulem J.E., Taylor K., Whitehead S., Barrett B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence." ;
 RN Nature 393:537-544 (1998).
 (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bhat W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLPTHIOCEROL
 CC SYNTHESIS.
 CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES (BY
 CC SIMILARITY). CONTAINS 2 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: Z74697; CAA38988.1; -;
 DR EMBL: AE007122; AAK47328.1; -;
 DR TIGR: MT3000; -;
 DR Tuberculest; RV2931; -;
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantone attach.
 DR Pfam: PF00109; ketoacyl-synt 1.
 DR Pfam: PF00550; pp-binding; 2.
 DR Pfam: PF00598; Acyl transfer; 1.
 DR Pfam: PF02801; ketoacyl-synt C; 1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE: PS00606; B KETOACYL SYNTHASE; 1.
 DR PROSITE: PS00075; ACP DOMAIN; 2.
 KW Multifunctional enzyme; Oxidoreductase; Transferase; NADP; Repeat;
 KW Phosphopantetheine; Complete proteome.
 FT DOMAIN 7 80 ACYL CARRIER (ACP) 1.
 FT NP BIND 1498 1503 ACYL CARRIER (ACP) 2.
 FT BINDING 62 62 NADP (POTENTIAL).
 FT ACT SITE 273 273 PHOSPHOPANTETHEINE (POTENTIAL).
 FT ACT SITE 720 720 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT BINDING 1796 1796 MALONYLTRANSFERASE (BY SIMILARITY).
 FT CONFLICT 624 624 D -> E (IN REF. 2).
 FT CONFLICT 877 877 R -> H (IN REF. 2).
 FT CONFLICT 1323 1323 G -> S (IN REF. 2).
 SQ SEQUENCE 1876 AA; 198834 MW; D9783DBD48792110 CRC64;

Query Match 19.1%; Score 81; DB 1; Length 1876;
 Best local similarity 34.1%; Pred. No. 0.44;
 Matches 15; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

DB 1789 PAHLGPDGWSGVELNRLSKATLRPLVTLIPHTTPAAVAARL 86
 FAEILGINSIMAMAIIRREAEQFVIEISATWLFMHPTVKLSAYL 1832

RESULT 9
 ID FAS_RAT STANDARD; PRT; 2505 AA.
 AC P12785; Q64717; O09187; O09190;

DT 01-OCT-1989 (rel. 12, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 15-JUL-1999 (rel. 38, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [Includes: EC 2.3.1.38; EC 2.3.1.39;
 DE EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89240686; PubMed=2717611;
 RA Amy C.M., Witkowski A., Naggett J., Williams B., Randhawa Z.,
 RA Smith S.;
 RT "Molecular cloning and sequencing of cDNAs encoding the entire rat
 RT fatty acid synthase." ;
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3114-3118 (1989).
 (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=93075999; PubMed=139331;
 RA Beck K.F., Schreglmann R., Stachopulos I., Klein H., Hoch J.,
 RA Schweizer M.;
 RT "The fatty acid synthase (FAS) gene and its promoter in Rattus
 RT norvegicus." ;
 RL DNA Seq. 2:359-366 (1992).
 (3)
 RP SEQUENCE OF 75-2505 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Mammary gland;
 RX MEDLINE=89128431; PubMed=2915923;
 RA Schweitzer M., Takebayashi K., Beck K.F., Schreglmann R.;
 RT "Rat mammary gland fatty acid synthase: localization of the
 RT constituent domains and two functional polyadenylation/termination
 RT signals in the cDNA." ;
 RL Nucleic Acids Res. 17:567-586 (1989).
 (4)
 RP SEQUENCE OF 2085-2505 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=88087240; PubMed=2891707;
 RA Naggett J., Witkowski A., Mikkelsen J., Smith S.;
 RT "Molecular cloning and sequencing of a cDNA encoding the thioesterase
 RT domain of the rat fatty acid synthetase." ;
 RL J. Biol. Chem. 263:1146-1150 (1988).
 (5)
 RP SEQUENCE OF 1921-2324 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=87246646; PubMed=3109907;
 RA Witkowski A., Naggett J., Mikkelsen J., Smith S.;
 RT "Molecular cloning and sequencing of a cDNA encoding the acyl carrier
 RT protein and its flanking domains in the mammalian fatty acid
 RT synthetase." ;
 RL Eur. J. Biochem. 165:601-606 (1987).
 CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxypalmitoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-

CC carrier protein] + oleate.
 CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M76767; AAA57219.1; -
 CC EMBL: X62888; CAA44679.1; -
 CC EMBL: X62889; CAA44680.1; -
 CC EMBL: X13415; CAA31780.1; -
 CC EMBL: X13527; CAA31882.1; -
 CC EMBL: J03514; AAA41144.1; -
 CC PIR: A30313; XTRTPA.
 CC InterPro: IPR001227; Ac transferase.
 CC InterPro: IPR002085; Adh zn family.
 CC InterPro: IPR000794; Ketoacyl-synt.
 CC InterPro: IPR003880; Pantine attach.
 CC InterPro: IPR000051; SAM_bind.
 CC InterPro: IPR001031; Thioesterase.
 CC Pfam: PF00107; adh_zinc; 1.
 CC Pfam: PF00550; pp-binding; 1.
 CC Pfam: PF00698; Acyl_transf; 1.
 CC Pfam: PF00975; Thioesterase; 1.
 CC Pfam: PF02801; ketoacyl-synt_C; 1.
 CC PROSITE: PS00012; PHOSPHOPANTETHEINE; 1.
 CC PROSITE: PS00606; B KETOACYL SYNTHASE; 1.
 CC PROSITE: PS50075; ACP DOMAIN; 1.
 CC Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 CC Hydrolyase; Oxidoreductase; Transferase; Lyase; NADP;
 CC Lyrioxal phosphate.
 CC DOMAIN 1 413
 CC FT 429 817 BETA-KETOACYL SYNTHASE.
 CC FT 1629 1857 ACYL AND MALONYL TRANSFERASES.
 CC FT 1858 2113 ENOYL REDUCTASE.
 CC FT 2118 2174 BETA-KETOACYL REDUCTASE.
 CC FT 2202 2505 ACT CARRIER (ACP).
 CC FT 161 161 THIOESTERASE.
 CC FT 581 581 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 CC FT 1665 1682 MALONYLTRANSFERASE (BY SIMILARITY).
 CC FT 1698 1698 NADP (ER).
 CC FT 1765 1780 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC FT 2151 2151 NADP (KR).
 CC FT 2302 2302 PHOSPHOPANTETHEINE (BY SIMILARITY).
 CC FT 2475 2475 THIOESTERASE (BY SIMILARITY).
 CC FT 878 878 BETA-HYDROXYACYL DEHYDRATASE (BY
 CC ACT_SITE
 CC FT 871 871 S -> P (IN REF. 3).
 CC FT 1967 1968 MV -> IL (IN REF. 5).
 CC FT 2085 2085 C -> P (IN REF. 4).
 CC FT 2106 2106 A -> V (IN REF. 1 AND 5).
 CC FT 2296 2296 Y -> H (IN REF. 1 AND 5).
 CC FT 2505 2505 AA: 272647 MM; 5810EC130373114 CRC64;
 CC SEQUENCE
 CC Query Match 19.1%; Score 81; DB 1; Length 2505;
 CC Best Local Similarity 31.2%; Pred. No. 0.63;
 CC Matches 24; Conservative 9; Mismatches 30; Indels 14; Gaps 1;
 CC
 CC QY 8 QOOSLLDLVRATMAVNDGNE-----RASAGSASAGAHGFGSWSG 53
 CC DB 2095 QPNAVSSFFLAKKVAHGDGRDLVKAHAIIGIRDLGINDSLADIGLDSWG 2154
 CC QY 54 VELRNRLSKATGRLPLV 70
 CC DB 2155 VEVKRIERHDLVLP 2171
 CC
 CC RESULT 10

DHBF BACSU
 ID DHBF BACSU STANDARD; PRT: 1278 AA.
 AC P45745;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable serine activating enzyme.
 GN DHBF.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxId=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertoro M.G., Bessières P., Bolotin A., Borchert S.,
 RA Boriss R., Boursier L., Brans A., Braun M., Briganti S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denicof F., Devine K.M., Dusterhoff A., Ehrlich S.D., Emerson P.T.,
 RA Ertlen K.D., Errington J., Fabre C., Ferrati E., Fougere D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Hatech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karmata D., Kasahara Y., Klaerr-Bianhard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Puig C., Purnelle B., Rapoport G., Ray M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose W., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccini E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terasaki P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vanier F., Vasseroiti A.,
 RA Viari A., Wambut R., Wedler E., Wedler H., Weltzenegger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasuno K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RL Nature 390:249-256 (1997).
 RN [2]
 RP SEQUENCE OF 1-95 FROM N.A.
 RC STRAIN=168 / Marburg;
 RX MEDLINE=96146538; PubMed=8550523;
 RA Rowland B.M., Taber H.W.;
 RT "Duplicate isochromatase synthase genes of Bacillus subtilis:
 RT regulation and involvement in the biosyntheses of menaquinone and
 RT 2,3-dihydroxybenzoate".
 RL J. Bacteriol. 178:854-861 (1996).
 CC -1- COPOLYMER: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTETHEINE (BY
 CC similarity)
 CC -1- PATHWAY: 2,3-dihydroxybenzoate biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z59120; CAB15186.1; -
 CC EMBL: U26444; AAC44634.1; -
 CC HSP: P14687; IAMU.

DR Subtilisin, BGI1243; dbp.
 DR InterPro; IPR000873; AMP-bind.
 DR InterPro; IPR001242; Condensatn.
 DR InterPro; IPR003880; Ppantne attach.
 DR Pfam; PF00501; AMP-binding; 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR Pfam; PF00668; Condensation; 2.
 DR PRINTS; PR00154; AMPBINDING.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; FALSE_NEG.
 DR PROSITE; PS00455; AMP BINDING; 1.
 DR PROSITE; PS50075; ACP DOMAIN; 1.
 DR Ligase; Phosphopantetheine; Complete proteome.
 FT DOMAIN 966 1033 ACP CARRIER (ACP).
 FT BINDING 996 996 PHOSPHOPANTHETHEINE (POTENTIAL).
 SQ SEQUENCE 1278 AA; 142068 MW; B44FBF3FCB08584 CRC64;

Query Match 18.5%; Score 78.5; DB 1; Length 1278;
 Best Local Similarity 28.8%; Pred. No. 0.55;
 Matches 23; Conservative 13; Mismatches 33; Indels 11; Gaps 2;

QY 7 TOGSLDLVRAHTMAVLNDGNERTASDAGPSAFHIGPDSVNGVELRNLSKATGL 66
 DB 964 TPBEELDLDF-ABVGL-----AVVGIDDSFELGSHLAARLMSKIREVMGA 1012
 QY 67 RLEFVTLIFDHTTPAAVAARL 86
 DB 1013 ELGIATLFDDEPTVAGLAARL 1032

RESULT 11
 ID_FAS_MOUSE STANDARD; PRT; 838 AA.

AC P19096;
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [includes: EC 2.3.1.38; EC 2.3.1.39;
 EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14]
 DE (Fragment).
 GN FASN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Liver;
 RX MEDLINE=89149781; PubMed=2920037;
 RA Paulaukis J.D., Sul H.S.;
 RT "Structure of mouse fatty acid synthase mRNA. Identification of the
 RT two NADPH binding sites";
 RT Biochem. Biophys. Res. Commun. 158:690-695(1989).
 CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACP CARRIER PROTEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) COA + N CO(2) + 2N NADP(+).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl]-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxypalmitoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl]-
 CC carrier protein] + oleate.

CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X13135; CAA1525.1; -.
 DR F1R; A32262; A32262.
 DR MGD; MG1:95485; Fasn.
 DR InterPro; IPR002085; Adh zn family.
 DR InterPro; IPR000794; Ketocacyl-synt.
 DR InterPro; IPR003880; Ppantne attach.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR Pfam; PF00975; Thioesterase; 1.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; PARTIAL.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
 DR PROSITE; PS50075; ACP DOMAIN; 1.
 DR Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 KW Hydrolyase; Oxidoreductase; Transferase; Lyase; NADP.
 FT NON TER 1 1
 FT DOMAIN 451 507 ACP CARRIER (ACP).
 FT BINDING 484 484 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT ACT_SITE 635 635 THIOESTERASE (BY SIMILARITY).
 FT ACT_SITE 808 808 THIOESTERASE (BY SIMILARITY).
 SQ SEQUENCE 838 AA; 91212 MW; 32C40E5C5A5B841 CRC64;

Query Match 18.4%; Score 78; DB 1; Length 838;
 Best Local Similarity 29.9%; Pred. No. 0.38;
 Matches 23; Conservative 11; Mismatches 29; Indels 14; Gaps 1;

QY 8 QOOSLLDLVRAHTMAVLNDGNE-----RTASDAGPSAFHIGPDSVNG 53
 DB 428 QPRAVLSFVLAKEKAVAHGDCGNDQRYKAVAHILGIDLAGINDSLADGLDLSWG 487

QY 54 VELRNLSKATGLRPV 70
 DB 488 VEVRLERHDLVLPW 504

RESULT 12
 ID_FAS_CHICK STANDARD; PRT; 2511 AA.

AC P12276;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [includes: EC 2.3.1.38; EC 2.3.1.39;
 EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN OR FAS.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-11.
 RC STRAIN=White Leghorn; TISSUE=Liver;
 RX MEDLINE=95031085; PubMed=7944406;
 RA Huang W.-Y., Chirlila S.S., Wakil S.J.;
 RT "Amino-terminal blocking group and sequence of the animal fatty acid
 RT synthase.";
 RL Arch. Biochem. Biophys. 314:45-49(1994).
 RN [2]
 RP SEQUENCE OF 75-1775 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89282777; PubMed=2734291;

RA Holzer K.P., Liu W., Hammes G.G.;
 RT "Molecular cloning and sequencing of chicken liver fatty acid
 RT synthase cDNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4387-4391(1989).
 RN [3]
 RN SEQUENCE OF 1568-2512 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=89139426; PubMed=2917973;
 RA Chirala S.S., Kasturi R., Pazirandeh M., Stollow D.T., Huang W.-Y.,
 Makil S.J.;
 RT "A novel cDNA extension procedure. Isolation of chicken fatty acid
 RT synthase cDNA clones.";
 RL J. Biol. Chem. 264:3750-3757(1989).
 RN [4]
 RN SEQUENCE OF 1752-2512 FROM N.A.
 RX MEDLINE=88320436; PubMed=2842766;
 RA Yuan Z., Liu W., Hammes G.G.;
 RT "Molecular cloning and sequencing of DNA complementary to chicken
 RT liver fatty acid synthase mRNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:6328-6331(1988).
 RN [5]
 RN SEQUENCE OF 2202-2512 FROM N.A.
 RX MEDLINE=89088152; PubMed=3207710;
 RA Kasturi R., Chirala S.S., Pazirandeh M., Makil S.J.;
 RT "Characterization of a genomic and cDNA clone coding for the
 RT cholesteryl domain and 3' noncoding region of the chicken liver
 RT fatty acid synthase gene.";
 RL Biochemistry 27:7778-7785(1988).
 RN [6]
 RN SEQUENCE OF 2121-2209.
 RX MEDLINE=89192401; PubMed=2648999;
 RA Huang W.-Y., Scoops J.K., Makil S.J.;
 RT "Complete amino acid sequence of chicken liver acyl carrier protein
 RT derived from the fatty acid synthase.";
 RL Arch. Biochem. Biophys. 270:92-98(1989).
 RN [7]
 RN SEQUENCE OF 2209-2508.
 RX STRAIN=White Leghorn;
 RA MEDLINE=89088151; PubMed=3207709;
 RA Yang C.-Y., Huang W.-Y., Chirala S.S., Makil S.J.;
 RT "Complete amino acid sequence of the cholesteryl domain of chicken
 RT liver fatty acid synthase.";
 RL Biochemistry 27:7773-7777(1988).
 RN [8]
 RN SEQUENCE OF 667-674 AND 1698-1709.
 RX MEDLINE=89323081; PubMed=2751995;
 RA Chang S.I., Hammes G.G.;
 RT "Amino acid sequences of pyridoxal 5'-phosphate binding sites and
 RT fluorescence resonance energy transfer in chicken liver fatty acid
 RT synthase.";
 RL Biochemistry 28:3781-3788(1989).
 CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) COA + N CO(2) + 2N NADP(+).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxypalmitoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-
 CC carrier protein] + oleate.
 CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC -1- CAUTION: REF.3 AND REF.5 SEQUENCES DIFFER FROM THAT SHOWN FROM

CC POSITION 2351 ONWARDS DUE TO A FRAMESHIFT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J04485; AAB46389.1; -
 CC EMBL: J03860; AAA48767.1; -
 CC EMBL: J02839; AAA82106.1; ALT_SEQ.
 CC PIR: A33918; XYCHPA.
 CC PIR: A33015; A32015.
 CC InterPro: IPR001227; Ac transferase.
 CC InterPro: IPR002085; Adh_zn_family.
 CC InterPro: IPR000794; Ketoacyl-synt.
 CC InterPro: IPR003880; Pantane attach.
 CC InterPro: IPR001031; Thioesterase.
 CC Pfam: PF00107; adh_zinc.1.
 CC Pfam: PF00109; ketoacyl-synt.1.
 CC Pfam: PF00550; pp-binding.1.
 CC Pfam: PF00698; Acyl_transf.1.
 CC Pfam: PF00975; Thioesterase.1.
 CC Pfam: PF02801; ketoacyl-synt C.1.
 CC PROSITE: PS00012; PHOSPHOPANTHETHEINE.1.
 CC PROSITE: PS00606; B_KETOACYL_SYNTHASE.1.
 CC PROSITE: PS00075; ACP DOMAIN.1.
 CC Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 CC Transferrase; Hydrolyase; Oxidoreductase; Lyase; NADP; Acetylation;
 CC Alternative splicing; Pyridoxal phosphate.
 CC KW TRANSFERASE; HYDROLASE; OXIDOREDUCTASE; LYASE; NADP; ACETYLATION;
 CC KX Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 CC Alternative splicing; Pyridoxal phosphate.
 CC FT INIT MET 0
 CC FT DOMAIN 1 7411
 CC FT DOMAIN 427 815
 CC FT DOMAIN 1638 1866
 CC FT DOMAIN 1867 2119
 CC FT DOMAIN 2124 2180
 CC FT DOMAIN 2209 2511
 CC FT MOD RES 1 1
 CC FT ACT_SITE 160 160
 CC FT ACT_SITE 579 579
 CC FT ACT_SITE 877 877
 CC FT NP BIND 1674 1691
 CC FT BINDING 1707 1707
 CC FT NP BIND 1888 1903
 CC FT BINDING 2157 2157
 CC FT ACT_SITE 2308 2308
 CC FT ACT_SITE 2481 2481
 CC FT VARSELIC 2348 2348
 CC FT CONFLICT 77 78
 CC FT CONFLICT 116 116
 CC FT CONFLICT 675 675
 CC FT CONFLICT 1169 1169
 CC FT CONFLICT 1178 1178
 CC FT CONFLICT 1191 1191
 CC FT CONFLICT 1198 1198
 CC FT CONFLICT 1286 1287
 CC FT CONFLICT 1372 1372
 CC FT CONFLICT 1533 1533
 CC FT CONFLICT 1577 1577
 CC FT CONFLICT 1685 1686
 CC FT CONFLICT 1732 1732
 CC FT CONFLICT 1745 1745
 CC FT CONFLICT 2511 AA; 274648 MM; 622039DNC81503F CRC64;
 CC SO SEQUENCE
 CC Query Match 17.3%; Score 73.5; DB 1; Length 2511;
 CC Best local Similarity 27.9%; Pred. No. 4,6;
 CC Matches 19; Conservative 12; Mismatches 24; Indels 13; Gaps 1;

OY 8 QOOSLLDVRATMAVLNDGNE-----RTASDAGPSAFHAGFDSVWGV 54
 DB 2102 QHPVWSSFFVLAKEVSKSGSQRDVVAHILGVRDVSLSMASSSLADGLDLSLMGV 2161
 OY 55 ELNRRLSK 62
 DB 2162 EVROTTLER 2169

RESULT 13

GALE_LACCA STANDARD; PRT; 331 AA.
 ID GALE_LACCA
 AC 084903;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-galactose 4-epimerase).
 OS Lactobacillus casei.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 CC Lactobacillus
 OX NCBI_TaxId=1582;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=64H;
 RX MEDLINE=98268934; PubMed=9603808;
 RA Bactenbrock K., Albert C.A.;
 RT "The gal genes for the Leloir pathway of Lactobacillus casei 64H.";
 RL Appl. Environ. Microbiol. 64:2013-2019(1998).
 CC -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -1- COPACTOR: NAD.
 CC -1- PATHWAY: Galactose metabolism; third step.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF05933; AAC19329.1; -
 DR HSBP; P09147; 1XEL.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR000205; NAD_binding.
 DR Pfam; PF01370; Epimerase; 1.
 DR TIGRFAMs; TIGR01179; GALE; 1.
 PT. IGEMERASE; NAD; Galactose metabolism.
 KM NP_BIND 2 33 NAD (POTENTIAL).
 SQ SEQUENCE 331 AA; 36338 MW; 420DBD3311EAVBP CRC64;

Query Match 17.1%; Score 72.5; DB 1; Length 331;
 Best Local Similarity 31.1%; Pred. No. 0.53; Mismatches 19; Conservative 13; Indels 9; Gaps 2;
 Matches 19; Conservative 13; Mismatches 19; Indels 9; Gaps 2;

OY 13 LLDVRAHTMAV.LNDGNERTASDAGPSAFHAGFDSVWGVLRNRLSKATGLRLPVT 71
 DB 226 VVDLADAHILALRYLDAGNKSASFNIGSAHGFSTL-----ELNAAKRTVGQETPAT 277
 OY 72 L 72
 DB 278 M 278

RESULT 14
 ID PURL_LACCA STANDARD; PRT; 740 AA.
 AC P35852;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphoribosylformylglycinamide synthase II (EC 6.3.5.3) (FGAM synthase II).
 GN PURL.
 OS Lactobacillus casei.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 CC Lactobacillus
 OX NCBI_TaxId=1582;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93012962; PubMed=1398079;
 RA Gu Z.-M., Martindale D.W., Lee B.H.;
 RT "Isolation and complete sequence of the purL gene encoding FGAM synthase II in Lactobacillus casei.";
 RL Gene 119:123-126(1992).
 FT (2)

RP REVISIONS.
 RX MEDLINE=94040790; PubMed=8224889;
 RA Gu Z.-M., Martindale D.W., Lee B.H.;
 RL Gene 133:147-147(1993).
 CC -1- CATALYTIC ACTIVITY: ATP + N(2)-(formyl-N(1)-(5-phospho-D-riboseyl)glycinamide + L-glutamine + H(2)O = ADP + phosphate + 2-(formamido)-N(1)-(5-phospho-D-riboseyl)acetamide + L-glutamate.
 CC -1- PATHWAY: De novo purine biosynthesis; fourth step.
 CC -1- SUBUNIT: HETERODIMER OF TWO SUBUNITS, PURQ AND PURL.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE FGAMS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M85265; AAC36947.1; -
 DR PIR; JCI290; JCI290.
 DR InterPro; IPR000728; AIRS-related.
 DR Pfam; PF00586; AIRS; 2.
 DR Pfam; PF02769; AIRS_C1; 2.
 KM Purine biosynthesis; Ligase; ATP-binding.
 NP_BIND 112 123 ATP (POTENTIAL).
 SQ SEQUENCE 740 AA; 79481 MW; 03238AB3CD0CBFF5 CRC64;

Query Match 17.0%; Score 72; DB 1; Length 740;
 Best Local Similarity 29.3%; Pred. No. 1.6; Mismatches 17; Conservative 10; Indels 25; Gaps 1;
 Matches 17; Conservative 10; Mismatches 17; Indels 25; Gaps 1;

OY 8 QOOSLLDVRATMAVLNDGNEERTASDAGPSAFHAGFDSVWGVLRNRLSKATG 65
 DB 618 QHOHFVLAKEVSKSGSQRDVVAHILGVRDVSLSMASSSLADGLDLSLMGV 669

RESULT 15
 ID LYS2_YEAST STANDARD; PRT; 1392 AA.
 AC P07762;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aminoadipate-semialdehyde dehydrogenase large subunit (EC 1.2.1.31)
 DE (Alpha-aminoacidate reductase) (Alpha-AR).
 GN LYS2 OR YBR115C OR YBR0910.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxId=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91192607; PubMed=2013406;
 RA Morris M.E., Jinks-Robertson S.;
 RT "Nucleotide sequence of the lys2 gene of Saccharomyces cerevisiae:

RT homology to Bacillus brevis tyrocidine synthetase 1.";
 RL Gene 98:141-145(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=8288C;
 RA MEDLINE=95208357; PubMed=7900426;
 RA Mannhaupt G., Stucka R., Ehnl S., Vetter I., Feldmann H.;
 RT "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
 RL Yeast 10:1363-1381(1994).
 RN [3]
 RP SEQUENCE OF 1-150 AND 1209-1392 FROM N.A.
 RX MEDLINE=87106859; PubMed=3542721;
 RA Fleig U.N., Fridmore R.D., Philippsen P.;
 RT "Construction of LYS2 cassettes for use in genetic manipulations of
 RT Saccharomyces cerevisiae.";
 RL Gene 46:237-245(1986).
 RN [4]
 RP SEQUENCE OF 1083-1392 FROM N.A.
 RX STRAIN=8288C;
 RA MEDLINE=92327848; PubMed=1626431;
 RA Mannhaupt G., Stucka R., Ehnl S., Vetter I., Feldmann H.;
 RT "Molecular analysis of yeast chromosome II between CMD1 and LYS2: the
 RT excision repair gene RAD52 located in this region belongs to a novel
 RT group of double-finger proteins.";
 RL Yeast 8:397-408(1992).
 RN [5]
 RP SEQUENCE OF 1-130 FROM N.A.
 RX STRAIN=8288C;
 RA MEDLINE=94039074; PubMed=7916691;
 RA Schaeff-Gerstenlachger I., Mannhaupt G., Vetter I., Zimmermann F.K.,
 RA Feldmann H.;
 RT "FKL2, a second transketolase gene of Saccharomyces cerevisiae.
 RT Cloning, sequence and deletion analysis of the gene.";
 RL Eur. J. Biochem. 217:487-492(1993).
 CC -1- FUNCTION: CATALYZES THE ACTIVATION OF ALPHA-AMINOADIPATE BY
 CC ATP-DEPENDENT ADENYLATION AND THE REDUCTION OF ACTIVATED
 CC ALPHA-AMINOADIPATE BY NADPH.
 CC CATALYTIC ACTIVITY: L-2-aminoadipate 6-semialdehyde + NADP(+) +
 CC H(2)O = L-2-aminoadipate + NADPH.
 CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTETHEINE
 CC (POTENTIAL).
 CC -1- PATHWAY: Lysine biosynthesis; sixth step.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M36287; AAA34747.1; -
 DR EMBL; X66247; CAA46975.1; -
 DR EMBL; X78993; CAA55617.1; -
 DR EMBL; Z35984; CAA85072.1; -
 DR EMBL; X73532; CAA51938.1; -
 DR PIR; J00448; YGBYAD.
 DR HSSP; P14687; LAMU.
 DR SGD; S0000319; LYS2.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR Pfam; PF00550; PP-binding; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE; PS00455; AMP BINDING; 1.
 DR PROSITE; PS00075; ACP DOMAIN; 1.
 KM Lysine biosynthesis; Oxidoreductase; NADP, Phosphopantetheine.
 FT DOMAIN 848 917 ACYL CARRIER (ACP).
 FT BINDING 880 880 PHOSPHOPANTETHEINE (POTENTIAL).

SQ SEQUENCE 1392 AA; 155345 MW; P0083A80BC6F7FB5 CRC64;
 Query Match 17.0%; Score 72; DB 1; Length 1392;
 Best Local Similarity 28.1%; Pred. No. 3.4;
 Matches 27; Conservative 14; Mismatches 37; Indels 18; Gaps 3;
 QY 6 PTOOSILLDLVRAHTMAVLND---GNERTASD-----AGPSASFAHLGFDG 50
 DB 824 PTPKO---LNLVAENTVSETDSQFTNVEREVDLWLSILPTKPASVSPDPSFFDLGGHS 880
 QY 51 VMGVEILRNLSKATGRLPVTLLIFDHTTPAAVAARL 86
 DB 881 ILATKMIPTLKKKLQVDLPGLGTLFKYPTIKAFAAEI 916

Search completed: June 17, 2003, 13:02:32
 Job time : 1.93786 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:55:32 ; Search time 1.92037 Seconds
(without alignments)
4305.183 Million cell updates/sec

Title: US-09-914-286-3_COPY_366_451
Perfect score: 424
Sequence: 1 LTALSPTQOOSLLDLVRAHMAVNDGNERFASDPGSAFAHLGPDVWGVELRNRL 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	199	46.9	4613	2 T17409	polyketide synthase
2	195	46.0	1862	2 T17411	polyketide synthase
3	193	45.5	3413	2 T17467	rifamycin polyketide
4	189	44.6	2100	2 T03223	probable polyketide
5	184	43.4	1763	2 T17465	rifamycin polyketide
6	184	43.4	5069	2 T17464	rifamycin polyketide
7	183	43.2	4735	2 T17463	rifamycin polyketide
8	182	42.9	3739	2 T17410	polyketide synthase
9	181	42.7	1762	2 T03222	probable polyketide
10	180	42.5	6260	2 T03228	polyketide synthase
11	180	42.5	10223	2 T30225	polyketide synthase
12	179	42.2	1616	2 G70668	polyketide synthase
13	178	42.0	8563	2 T30226	polyketide synthase
14	176.5	41.6	7576	2 T17428	polyketide synthase
15	176	41.5	1728	2 T17466	rifamycin polyketide
16	174	41.0	2103	2 G86925	probable 6-deoxyer
17	173	40.8	3491	2 T43231	probable polyketide
18	169	39.9	2126	2 T70621	polyketide synthase
19	168	39.6	1198	2 T28678	polyketide synthase
20	168	39.6	4151	2 G70944	probable polyketide
21	168	39.6	6420	2 T30283	polyketide synthase
22	166	39.2	3519	2 S43048	polyketide synthase
23	162.5	38.3	502	2 A70985	probable polyketide
24	161	38.0	3573	2 S23070	erythronolide synth
25	160.5	37.9	1017	2 B70985	probable polyketide
26	155	36.6	2124	2 T28658	polyketide synthase
27	150	35.4	1402	2 D70634	probable polyketide
28	138	32.5	3172	2 S22012	erythronolide synth
29	138	32.5	3178	2 S13595	6-deoxyerythronol

30	137	32.3	1346	2 T17412	polyketide synthase
31	135	31.8	2723	2 T03221	probable polyketide
32	133	31.4	1937	2 T03224	probable polyketide
33	109	25.7	2201	2 S73014	polyketide synthase
34	104	24.5	707	2 F86925	probable acyl-CoA
35	104	24.5	2188	2 A70984	probable polyketide
36	102	24.1	705	2 A70669	probable acid-CoA
37	99	23.3	1538	2 E70874	probable ppsb prot
38	96.5	22.8	89	2 S18958	fix3-6 protein -
39	96	22.6	1540	2 H87203	polyketide synthase
40	94.5	22.3	130	2 S73018	hypothetical prote
41	91.5	21.6	622	2 A61197	6-methylsalicylic
42	90.5	21.3	1774	2 S13178	6-methylsalicylic
43	90	21.2	8243	2 T31307	type I fatty acid
44	88.5	20.9	15281	2 S41309	cyclosporin synth
45	87	20.5	1004	2 H70673	probable mbtd prot

ALIGNMENTS

RESULT 1
T17409
polyketide synthase type I - Streptomyces venezuelae
C/Species: Streptomyces venezuelae
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #ext_change 17-Nov-2000
C/Accession: T17409
R/Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
A/Title: A gene cluster for macrolide antibiotic biosynthesis in streptomyces venezuelae
A/Reference number: Z18773; MUID:98445333; PMID:9770448
A/Accession: T17409
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-4613 <XUE>
A/Cross-references: EMBL:AF079138; NID:G3808326; PID:G3800834; PIDN:AA669329.1
C/Genetics:
A/Genes: pikA
C/Superfamily: acyl carrier protein homology
C/Keywords: antibiotic biosynthesis; carrier protein
F/1010-1081/Domain: acyl carrier protein homology <ACPI>
F/2495-2566/Domain: acyl carrier protein homology <ACP2>
F/4407-4478/Domain: acyl carrier protein homology <ACP3>

Query Match
Best Local Similarity 46.9%; Score 199; DB 2; Length 4613;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;
Db 2485 LAAAGGERTELLGIVRQAAV-----RMSPEVADRAFKDGFSLAGVELRNRL 2540

QY 1 LTALSPTQOOSLLDLVRAHMAVNDGNERFASDPGSAFAHLGPDVWGVELRNRL 60
DB 2485 LAAAGGERTELLGIVRQAAV-----RMSPEVADRAFKDGFSLAGVELRNRL 2540

QY 61 SKATGLRPVTLIFDHTTPAAVAARL 86
DB 2541 TRATGQLRPVTLIFDHTTPAAVAARL 2566

RESULT 2
T17411
polyketide synthase III - Streptomyces venezuelae
C/Species: Streptomyces venezuelae
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #ext_change 03-Nov-2000
C/Accession: T17411
R/Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
A/Title: A gene cluster for macrolide antibiotic biosynthesis in streptomyces venezuelae
A/Reference number: Z18773; MUID:98445333; PMID:9770448
A/Accession: T17411
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1562 <XUE>
A/Cross-references: EMBL:AF079138; NID:G3808326; PID:G3800836; PIDN:AA669331.1
C/Genetics:

Query Match 42.5%; Score 180; DB 2; Length 6260;
 Best Local Similarity 48.8%; Pred. No. 2,4e-11;
 Matches 42; Conservative 11; Mismatches 29; Indels 4; Gaps 1;

QY 1 LTLSPFQOOSLLDVRATMAVNDGNERFASDGPASPAHLGFDVSGVELRNL 60
 DB 3113 LALAPAPERDALKLVKVRSDAATLV---GHADASTIPAAAFKIDISLTVAVELRNL 3168
 61 SKATGLRLPTVTLFDHTTPAAVAARL 86
 DB 3169 AKATGLRLPTVTFDPTPALAARL 3194

RESULT 11
 T30225
 polyketide synthase - Streptomyces hygroscopicus
 C/Species: Streptomyces hygroscopicus
 C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
 C/Accession: T30225
 R/Particlo, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Staun
 Gene 169, 9-16, 1996
 A/Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hy
 A/Reference number: Z20782; MUID:96186896; PMID:8635756
 A/Accession: T30225
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-10223 <APA>
 A/Cross-references: EMBL:X86780; NID:g987088; PID:g987099; PIDN:CAA60459.1
 C/Genetics:
 A/Genes: rapB
 C/Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protei
 C/Keywords: carrier protein
 F/54-449/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
 F/1533-1604/Domain: acyl carrier protein homology <ACP1>
 F/1647-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
 F/2135-2409/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
 F/3172-3243/Domain: acyl carrier protein homology <ACP2>
 F/3287-3681/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
 F/3778-4052/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
 F/5160-5231/Domain: acyl carrier protein homology <ACP3>
 F/5275-5657/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
 F/6776-6847/Domain: acyl carrier protein homology <ACP4>
 F/6891-7285/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS5>
 F/8411-8482/Domain: acyl carrier protein homology <ACP5>
 F/5526-8921/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS6>
 F/9012-9285/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
 F/1066-10140/Domain: acyl carrier protein homology <ACP>

Query Match 42.5%; Score 180; DB 2; Length 10223;
 Best Local Similarity 48.8%; Pred. No. 4.3e-11;
 Matches 42; Conservative 11; Mismatches 29; Indels 4; Gaps 1;

QY 1 LTLSPFQOOSLLDVRATMAVNDGNERFASDGPASPAHLGFDVSGVELRNL 60
 DB 5150 LALAPAPERDALKLVKVRSDAATLV---GHADASTIPAAAFKIDISLTVAVELRNL 5205
 61 SKATGLRLPTVTLFDHTTPAAVAARL 86
 DB 5206 TKATGLRLPTVTFDPTPALAARL 5231

RESULT 12
 G70668
 polyketide synthase pksI - Mycobacterium tuberculosis (strain H37Rv)
 N/Alternate names: polyketide synthase pks02c
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C/Accession: G70668; S73075
 R/Coile, S.T.; Brosch, R.; Parish, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 ; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A/authors: Sgaree, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; MUID:98295987; PMID:9634230
 A/Accession: G70668
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-1616 <COL>
 A/Cross-references: GB:Z83858; GB:AL123456; NID:g3261675; PIDN:CAB06103.1; PID:g1781167
 A/Experimental source: strain H37Rv
 R/Smith, D.R.; Robison, K.
 submitted to the EMBL Data Library, September 1994
 A/Description: Mycobacterium tuberculosis cosmid tbc2.
 A/Reference number: S73053
 A/Accession: S73075
 A/Molecule type: DNA
 A/Residues: 'MMGARMVVARCLSTEL', 48-917, 'RSEARSRRGCGMGRAPPAPPIGPPIG', 'EMPRIPVBPAGPVT
 A/Cross-references: EMBL:U00024; NID:g560506; PIDN:AAA50928.1; PID:g560507
 C/Genetics:
 A/Genes: pksI
 C/Superfamily: Mycobacterium tuberculosis probable polyketide synthase pksI; acyl carrie
 C/Keywords: carrier protein
 F/84-362/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
 F/1228-1410/Domain: short-chain alcohol dehydrogenase homology <SADH>
 F/1514-1585/Domain: acyl carrier protein homology <ACP1>

Query Match 42.2%; Score 179; DB 2; Length 1616;
 Best Local Similarity 50.6%; Pred. No. 6.1e-12;
 Matches 42; Conservative 6; Mismatches 31; Indels 4; Gaps 1;

QY 1 LTLSPFQOOSLLDVRATMAVNDGNERFASDGPASPAHLGFDVSGVELRNL 60
 DB 1504 LHLADLEQLVGLVGLQAAATLV---GRSABEDVDTEFGDGLFSLTVAVELRNL 1559
 61 SKATGLRLPTVTLFDHTTPAAVA 83
 DB 1560 KATGLRLPTVTFDPTPALAARL 1582

RESULT 13
 T30226
 polyketide synthase - Streptomyces hygroscopicus
 C/Species: Streptomyces hygroscopicus
 C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 01-Dec-2000
 C/Accession: T30226
 R/Particlo, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Staun
 Gene 169, 9-16, 1996
 A/Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hy
 A/Reference number: Z20782; MUID:96186896; PMID:8635756
 A/Accession: T30226
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-8563 <APA>
 A/Cross-references: EMBL:X86780; NID:g987088; PID:g987100; PIDN:CAA60460.1
 C/Genetics:
 A/Genes: rapA
 C/Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase
 C/Keywords: carrier protein; phosphopantetheine; phosphoprotein
 F/54-503/Domain: acetate-CoA ligase homology <ACL>
 F/1329-1724/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
 F/1817-2091/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
 F/3199-3270/Domain: acyl carrier protein homology <ACP1>
 F/3314-3706/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
 F/4787-4858/Domain: acyl carrier protein homology <ACP2>
 F/4902-5293/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
 F/5386-5659/Domain: acyl carrier protein homology <ACP3>
 F/5760-6831/Domain: acyl carrier protein homology <ACP4>
 F/6875-7269/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
 F/7362-7636/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
 F/8412-8483/Domain: acyl carrier protein homology <ACP4>
 F/8822/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 42.0%; Score 178; DB 2; Length 8563;
 Best Local Similarity 48.8%; Pred. No. 5.9e-11;

```
A;Residues: 1-1728 <SCH>  
A;Cross-references: EMBL:AJ223012; NID:e1227119; PID:e1227123; PIDD:CAA11038.1  
C;Experimental source: strain LSG A3136  
C;Superfamily: Streptomycetes hygroscopticus probable polyketide synthase module 4; 3-Oxoacetyl-CoA ligase  
homology; [acyl-carrier-protein] S-malonyltransferase homology  
C;Keywords: carrier protein  
F;1573-1644/Domain: acyl carrier protein homology <ACP>
```

Query Match 41.5%; Score 176; DB 2; Length 1728;
Best Local Similarity 44.2%; Pred. No. 1.5e-11;
Matches 38; Conservative 14; Mismatches 30; Indels 4; Gaps 1;

Oy 1 LTALSPTQOQSLLDLPAAHTMAVINDGNETASDAGPSASFALGFDSVMGVLEARNRL 60
 | :::||:||||||| | | | | | | | | |
Db 1563 LAGLGAPQEALLVDLVRGQAAYLGHAGPDAVRAD---TAFKDGFDLSLTSDLRNL 1618
 | :::||:||||||| | | | | | | | | |

Oy 61 SKATGLRLPVTLIFDHTTPPAVAARL 86
 | :::||:||||||| | | | | | | | | |
Db 1619 RESTGLKPATLAEDYPPPLVLARNL 1644

Search completed: June 17, 2003, 13:12:19
Job time : 1.92037 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Gen Ltd.

OM protein - protein search, using SW model

Run on: June 17, 2003, 13:08:02 ; Search time 2.67959 Seconds
(without alignments)
3431.399 Million cell updates/sec

Title: US-09-914-286-3_COPY_366_451

Perfect score: 424
Sequence: 1 LTALSPPTQOQSLLDLVRAH.....RUPVLIFDHTTAAVAAVL 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubppa/PTCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB pep.*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB pep.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	199	46.9	4551	US-09-793-708-1	Sequence 1, App1
2	199	46.9	4613	US-09-860-846-31	Sequence 31, App1
3	199	46.9	4613	US-09-988-3848-31	Sequence 31, App1
4	199	46.9	4613	US-09-836-821-31	Sequence 31, App1
5	199	46.9	4613	US-09-861-289-31	Sequence 31, App1
6	199	46.9	11877	US-09-860-846-6	Sequence 6, App1
7	199	46.9	11877	US-09-836-821-6	Sequence 6, App1
8	199	46.9	11877	US-09-861-289-6	Sequence 6, App1
9	199	46.9	12199	US-09-988-3848-6	Sequence 6, App1
10	199	46.9	4150	US-09-808-880-2	Sequence 2, App1
11	199	46.0	1562	US-09-860-846-35	Sequence 35, App1
12	199	46.0	1562	US-09-988-3848-35	Sequence 35, App1
13	199	46.0	1562	US-09-836-821-35	Sequence 35, App1
14	199	46.0	1562	US-09-793-708-35	Sequence 35, App1
15	199	46.0	1562	US-09-861-289-35	Sequence 35, App1
16	199	43.2	5215	US-09-860-846-2	Sequence 2, App1
17	199	43.2	5215	US-09-988-3848-2	Sequence 2, App1
18	199	43.2	5215	US-09-836-821-2	Sequence 2, App1
19	199	43.2	5215	US-09-861-289-2	Sequence 2, App1

20	182	42.9	3739	US-09-860-846-33	Sequence 33, App1
21	182	42.9	3739	US-09-988-3848-33	Sequence 33, App1
22	182	42.9	3739	US-09-836-821-33	Sequence 33, App1
23	182	42.9	3739	US-09-793-708-33	Sequence 33, App1
24	182	42.9	3739	US-09-861-289-33	Sequence 33, App1
25	179	42.2	1616	US-09-712-363-262	Sequence 262, App1
26	166.5	39.3	3816	US-09-808-880-3	Sequence 3, App1
27	166	39.2	3519	US-09-808-880-4	Sequence 4, App1
28	150	35.4	1402	US-09-712-363-156	Sequence 156, App1
29	139	32.8	1346	US-09-793-708-4	Sequence 4, App1
30	137	32.3	1346	US-09-860-846-37	Sequence 37, App1
31	137	32.3	1346	US-09-988-3848-37	Sequence 37, App1
32	137	32.3	1346	US-09-836-821-37	Sequence 37, App1
33	137	32.3	1346	US-09-861-289-37	Sequence 37, App1
34	130	30.7	3798	US-10-014-717-6	Sequence 6, App1
35	117.5	27.7	1832	US-10-014-717-4	Sequence 4, App1
36	115.5	27.2	7257	US-10-014-717-5	Sequence 5, App1
37	115	27.1	352	US-09-073-009-14	Sequence 14, App1
38	115	27.1	352	US-09-023-568-14	Sequence 14, App1
39	115	27.1	352	US-09-793-306-14	Sequence 14, App1
40	103.5	24.4	2439	US-10-014-717-7	Sequence 7, App1
41	92.5	21.8	1421	US-10-014-717-2	Sequence 2, App1
42	83	19.6	4899	US-09-976-059-15	Sequence 15, App1
43	82	19.3	4399	US-09-976-059-14	Sequence 14, App1
44	79	18.6	1812	US-09-775-938A-38	Sequence 38, App1
45	78.5	18.5	90	US-09-976-059-12	Sequence 12, App1

ALIGNMENTS

RESULT 1
US-09-793-708-1
Sequence 1, Application US/09793708
Publication No. US20030104597A1
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002121
CURRENT APPLICATION NUMBER: US/09/793,708
CURRENT FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: US 09/657,440
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: US 09/320,878
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: US 09/141,908
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: US 09/073,538
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: US 08/846,247
PRIOR FILING DATE: 1997-04-30
PRIOR APPLICATION NUMBER: US 60/134,990
PRIOR FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 4551
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-793-708-1

Query Match 46.9%; Score 199; DB 9; Length 4551;

Best Local Similarity 48.8%; Pred. No. 9, 6e-16;

Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

QY 1 LTALSPPTQOQSLLDLVRAHVTAVNDGNETTASDAGSASFALGPDVWGVCLRNRL 60
DB 2423 LAAAPGERTETLLGLVRAQAALV-----RMRSPDVAAADRAFKOIGDFSLAGVELRNRL 2478

QY 61 SKATGLRLPVTLLIFDHTTPAAVAARL 86
Db 2479 TRATGQLPATLVFDPHTPLALVSL 2504

RESULT 2

US-09-860-846-31
; Sequence 31, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860.846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105.537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRF
; ORGANISM: Streptomyces venezuelae
US-09-860-846-31

Query Match 46.9%; Score 199; DB 9; Length 4613;
Best Local Similarity 48.8%; Pred. No. 9.8e-16;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

QY 1 LTAALPTQOOSLLDLVRAHTMAVLNDGNERITASDAGPSAFHIGFDSVWGVELRNRL 60
Db 2485 LAAAPGERTETLLGLVRAQAAVL---RMSPEVDVAADRAFKDIGDSLAGEVLRNRL 2540
QY 61 SKATGLRLPVTLLIFDHTTPAAVAARL 86
Db 2541 TRATGQLPATLVFDPHTPLALVSL 2566

RESULT 3

US-09-988-384B-31
; Sequence 31, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988.384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105.537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRF
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-31

Query Match 46.9%; Score 199; DB 9; Length 4613;
Best Local Similarity 48.8%; Pred. No. 9.8e-16;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

QY 1 LTAALPTQOOSLLDLVRAHTMAVLNDGNERITASDAGPSAFHIGFDSVWGVELRNRL 60
Db 2485 LAAAPGERTETLLGLVRAQAAVL---RMSPEVDVAADRAFKDIGDSLAGEVLRNRL 2540

QY 61 SKATGLRLPVTLLIFDHTTPAAVAARL 86
Db 2541 TRATGQLPATLVFDPHTPLALVSL 2566

RESULT 4

US-09-836-821-31
; Sequence 31, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836.821
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105.537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRF
; ORGANISM: Streptomyces venezuelae
US-09-836-821-31

Query Match 46.9%; Score 199; DB 9; Length 4613;
Best Local Similarity 48.8%; Pred. No. 9.8e-16;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

QY 1 LTAALPTQOOSLLDLVRAHTMAVLNDGNERITASDAGPSAFHIGFDSVWGVELRNRL 60
Db 2485 LAAAPGERTETLLGLVRAQAAVL---RMSPEVDVAADRAFKDIGDSLAGEVLRNRL 2540
QY 61 SKATGLRLPVTLLIFDHTTPAAVAARL 86
Db 2541 TRATGQLPATLVFDPHTPLALVSL 2566

RESULT 5

US-09-861-289-31
; Sequence 31, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861.289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105.537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRF
; ORGANISM: Streptomyces venezuelae
US-09-861-289-31

Query Match 46.9%; Score 199; DB 10; Length 4613;
Best Local Similarity 48.8%; Pred. No. 9.8e-16;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

QY 1 LTAALPTQOOSLLDLVRAHTMAVLNDGNERITASDAGPSAFHIGFDSVWGVELRNRL 60
Db 2485 LAAAPGERTETLLGLVRAQAAVL---RMSPEVDVAADRAFKDIGDSLAGEVLRNRL 2540
QY 61 SKATGLRLPVTLLIFDHTTPAAVAARL 86

Db 2541 TRATGQLPATLVFDHPTPLALVSL 2566

RESULT 6

US-09-860-846-6
Sequence 6, Application US/09860846
Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/860,846
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-860-846-6

Query Match 46.9%; Score 199; DB 9; Length 11877;
Best Local Similarity 48.8%; Pred. No. 3.5e-15;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

Qy 1 LTALSPTQOOSLLDLVRAHTMAVLNDGNERITASDAGPSAFHIGFDSVWGVELRNRL 60
Db 2821 LAAAPRGERTTEILGLVRAQAAVL-----RMSPEVDAAADRAFKDIGFSLAGVELRNRL 2876

Qy 61 SKATGLRLPVTLLIFDHTTPAAVAARL 86
Db 2877 TRATGQLPATLVFDHPTPLALVSL 2902

RESULT 7

US-09-836-821-6
Sequence 6, Application US/09836821
Publication No. US20030087405A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/836,821
CURRENT FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-836-821-6

Query Match 46.9%; Score 199; DB 9; Length 11877;
Best Local Similarity 48.8%; Pred. No. 3.5e-15;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

Qy 1 LTALSPTQOOSLLDLVRAHTMAVLNDGNERITASDAGPSAFHIGFDSVWGVELRNRL 60
Db 2821 LAAAPRGERTTEILGLVRAQAAVL-----RMSPEVDAAADRAFKDIGFSLAGVELRNRL 2876

Qy 61 SKATGLRLPVTLLIFDHTTPAAVAARL 86

Db 2877 TRATGQLPATLVFDHPTPLALVSL 2902

RESULT 8

US-09-861-289-6
Sequence 6, Application US/09861289
Patent No. US20020110897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/861,289
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-861-289-6

Query Match 46.9%; Score 199; DB 10; Length 11877;
Best Local Similarity 48.8%; Pred. No. 3.5e-15;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

Qy 1 LTALSPTQOOSLLDLVRAHTMAVLNDGNERITASDAGPSAFHIGFDSVWGVELRNRL 60
Db 2821 LAAAPRGERTTEILGLVRAQAAVL-----RMSPEVDAAADRAFKDIGFSLAGVELRNRL 2876

Qy 61 SKATGLRLPVTLLIFDHTTPAAVAARL 86
Db 2877 TRATGQLPATLVFDHPTPLALVSL 2902

RESULT 9

US-09-988-384B-6
Sequence 6, Application US/09988384B
Publication No. US20030073824A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.536US1
CURRENT APPLICATION NUMBER: US/09/988,384B
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: PCT/US99/14398
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 53
SEQ ID NO 6
LENGTH: 12199
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-988-384B-6

Query Match 46.9%; Score 199; DB 9; Length 12199;
Best Local Similarity 48.8%; Pred. No. 3.6e-15;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

Qy 1 LTALSPTQOOSLLDLVRAHTMAVLNDGNERITASDAGPSAFHIGFDSVWGVELRNRL 60
Db 3143 LAAAPRGERTTEILGLVRAQAAVL-----RMSPEVDAAADRAFKDIGFSLAGVELRNRL 3198

Qy 61 SKATGLRLPVTLLIFDHTTPAAVAARL 86

Db 13199 TRATGQLPATVFDHPPTPLATVSL 3224

RESULT 10

US-09-808-880-2
; Sequence 2, Application US/0980880
; Publication No. US20030027287A1
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/808,880
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US/09/428,517
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/120,254
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/106,100
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-808-880-2

Query Match 46.2%; Score 196; DB 9; Length 4150;
Best Local Similarity 50.0%; Pred. No. 2.1e-15;
Matches 43; Conservative 10; Mismatches 29; Indels 4; Gaps 1;

QY 1 LTALPTQOOSLLDLVRAHTMAVLNDGNERITASDAGSPASFAHIGFDSVNGVELRNRL 60
Db 2441 LAGLTVAEDEERLLVQVVRGAAYVLGHSGAEAV----FDRAKDGLGFSLSVLEARNL 2496

QY 61 SKATGRLPVTLLIFDHTTPAAVAARL 86
Db 2497 NTATGRLPVTAVFDYARPAALAGHL 2522

RESULT 11

US-09-860-846-35
; Sequence 35, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 1562
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-35

Query Match 46.0%; Score 195; DB 9; Length 1562;
Best Local Similarity 48.8%; Pred. No. 7.4e-16;
Matches 42; Conservative 11; Mismatches 29; Indels 4; Gaps 1;

QY 1 LTALPTQOOSLLDLVRAHTMAVLNDGNERITASDAGSPASFAHIGFDSVNGVELRNRL 60

Db 1393 ITALPEPERPALTLVTRHAAAVLGHSSPDVA----PGRATFELGFSLSLAVQLRNRL 1448

QY 61 SKATGRLPVTLLIFDHTTPAAVAARL 86
Db 1449 STVVGNRPLPATVFDHPPTPALAARL 1474

RESULT 12

US-09-988-384B-35
; Sequence 35, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 35
; LENGTH: 1562
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-35

Query Match 46.0%; Score 195; DB 9; Length 1562;
Best Local Similarity 48.8%; Pred. No. 7.4e-16;
Matches 42; Conservative 11; Mismatches 29; Indels 4; Gaps 1;

QY 1 LTALPTQOOSLLDLVRAHTMAVLNDGNERITASDAGSPASFAHIGFDSVNGVELRNRL 60
Db 1393 ITALPEPERPALTLVTRHAAAVLGHSSPDVA----PGRATFELGFSLSLAVQLRNRL 1448

QY 61 SKATGRLPVTLLIFDHTTPAAVAARL 86
Db 1449 STVVGNRPLPATVFDHPPTPALAARL 1474

RESULT 13

US-09-836-821-35
; Sequence 35, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 1562
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-836-821-35

Query Match 46.0%; Score 195; DB 9; Length 1562;
Best Local Similarity 48.8%; Pred. No. 7.4e-16;
Matches 42; Conservative 11; Mismatches 29; Indels 4; Gaps 1;

; SOFTWARE: FastSeq for Windows Version 3.0

Search completed: June 17, 2003, 13:46:40
Job time : 3.67959 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:54:32 ; Search time 1.5631 Seconds
(without alignments)
1618.821 Million cell updates/sec

Title: US-09-914-286-3_COPY_366_451
Perfect score: 424
Sequence: 1 LTALSPTOQOSLLDLVRAH.....RLPVTLPDHTTPAAVAARL 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/1aa/5A COMB.pep.*
2: /cgn2_6/prodata/1/1aa/5B COMB.pep.*
3: /cgn2_6/prodata/1/1aa/5A COMB.pep.*
4: /cgn2_6/prodata/1/1aa/5B COMB.pep.*
5: /cgn2_6/prodata/1/1aa/5A COMB.pep.*
6: /cgn2_6/prodata/1/1aa/5B COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	199	46.9	4551	US-09-320-878-1	Sequence 1, App1
2	199	46.9	4613	US-09-105-537-31	Sequence 31, App1
3	199	46.9	4928	US-09-036-987A-5	Sequence 5, App1
4	199	46.9	4928	US-09-370-700-5	Sequence 5, App1
5	199	46.9	11877	US-09-105-537-6	Sequence 2, App1
6	196	46.2	4150	US-09-428-517-2	Sequence 2, App1
7	195	46.0	1562	US-09-320-878-3	Sequence 3, App1
8	195	46.0	1562	US-09-105-537-35	Sequence 35, App1
9	192	45.3	2152	US-09-036-987A-3	Sequence 3, App1
10	192	45.3	2152	US-09-370-700-3	Sequence 3, App1
11	190	44.8	3724	US-08-804-227C-10	Sequence 10, App1
12	190	44.8	3724	US-08-804-198-4	Sequence 4, App1
13	190	44.8	3729	US-08-804-227C-4	Sequence 4, App1
14	183	43.2	5215	US-09-105-537-2	Sequence 2, App1
15	182	42.9	3739	US-09-320-878-2	Sequence 2, App1
16	182	42.9	3739	US-09-105-537-33	Sequence 33, App1
17	181	42.7	3170	US-09-036-987A-4	Sequence 4, App1
18	181	42.7	3170	US-09-370-700-4	Sequence 4, App1
19	180	42.5	185	US-09-154-083-14	Sequence 14, App1
20	178	42.0	1611	US-08-804-227C-5	Sequence 5, App1
21	177	41.7	502	US-09-413-814-106	Sequence 106, App1
22	176.5	41.6	2595	US-09-036-987A-2	Sequence 2, App1
23	176.5	41.6	2595	US-09-370-700-2	Sequence 2, App1
24	176	41.5	1996	US-08-804-227C-9	Sequence 9, App1
25	176	41.5	1996	US-08-804-198-3	Sequence 3, App1
26	174	41.0	3567	US-07-642-734C-4	Sequence 4, App1
27	174	41.0	3567	US-08-439-009A-4	Sequence 4, App1

28	173	40.8	3491	2	US-07-642-734C-2	Sequence 2, App1
29	173	40.8	3491	3	US-08-439-009A-2	Sequence 2, App1
30	171	40.3	5588	4	US-09-036-987A-6	Sequence 6, App1
31	171	40.3	5588	4	US-09-370-700-6	Sequence 6, App1
32	168	39.6	4472	2	US-08-804-227C-2	Sequence 2, App1
33	168	39.6	4455	2	US-08-804-227C-14	Sequence 14, App1
34	167	39.4	4550	2	US-08-804-198-2	Sequence 8, App1
35	167	39.4	4550	2	US-08-804-227C-8	Sequence 8, App1
36	166.5	39.3	3816	4	US-09-428-517-3	Sequence 4, App1
37	166	39.2	3519	4	US-09-428-517-4	Sequence 4, App1
38	164	38.7	1864	2	US-08-804-227C-3	Sequence 3, App1
39	162	38.2	1841	2	US-08-804-227C-6	Sequence 6, App1
40	159	37.5	1580	2	US-08-804-227C-11	Sequence 11, App1
41	159	37.5	1580	2	US-08-804-198-5	Sequence 5, App1
42	155	36.6	1891	2	US-08-804-227C-12	Sequence 12, App1
43	155	36.6	1891	2	US-08-804-198-6	Sequence 6, App1
44	152	35.8	5087	4	US-09-144-085-1	Sequence 1, App1
45	147	34.7	6095	4	US-09-144-085-2	Sequence 2, App1

ALIGNMENTS

```
RESULT 1
US-09-320-878-1
; Sequence 1, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OR INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; EARLIER FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4551
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-320-878-1

Query Match      46.9%; Score 199; DB 3; Length 4551;
Best Local Similarity 48.8%; Pred. No. 7.7e-17;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

QY 1 LTALSPTOQOSLLDLVRAHNTMAVINDGNETASDAGSPASFAHLGFPDVGVELRNRL 60
DB 2423 LAAABGERTETLLGLVRAQAANL---RMRSPEDVAADRAFKDIGFSLGVELRNRL 2478
QY 61 SKATGLRLPVTLPDHTTPAAVAARL 86
DB 2479 TRATGLQLPATLVDPHTPLALVSL 2504

RESULT 2
US-09-105-537-31
; Sequence 31, Application US/09105537A
```

```

; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ. ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-31

```

```

Query Match      46.9%; Score 199; DB 4; Length 4613;
Best Local Similarity 48.8%; Pred. No. 7.8e-17;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

```

```

QY      1 LTLSPFOOQSLLDLVRAHTMAVNDGNETASDAGPSAFHIGFDSVWGVELRNL 60
          |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
          2485 LAAPAEERTEIILGLVRAQAALV---RMRSPEVDVAADRAFKDIGFDSLAVELRNL 2540
          :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY      61 SKATGLRLPVTLLFDHTTPAAVAARL 86
          :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      2541 TRATGDLPLVFDHTPLALVSL 2566

```

RESULT 3

```

US-09-036-987A-5
; Sequence 5, Application US/09036987A
; Patent No. 6143526

```

```

; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merilo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4928 amino acids

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-036-987A-5

```

```

Query Match      46.9%; Score 199; DB 4; Length 4928;
Best Local Similarity 50.0%; Pred. No. 8.6e-17;
Matches 43; Conservative 13; Mismatches 26; Indels 4; Gaps 1;

```

```

QY      1 LTLSPFOOQSLLDLVRAHTMAVNDGNETASDAGPSAFHIGFDSVWGVELRNL 60
          |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
          4762 LAALPEARGHLLLEVLAEFTASTLGH-----SAAVQPDRTFAELGFDSTLAVELRNL 4817
          :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY      61 SKATGLRLPVTLLFDHTTPAAVAARL 86
          :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      4818 NAVTGLRLPVTLLFDHTPLALSEQ 4843

```

RESULT 4

```

US-09-370-700-5
; Sequence 5, Application US/09370700
; Patent No. 6274350

```

```

; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; EARLIER APPLICATION NUMBER: US 09/36987
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4928
; TYPE: PRT
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-5

```

```

Query Match      46.9%; Score 199; DB 4; Length 4928;
Best Local Similarity 50.0%; Pred. No. 8.6e-17;
Matches 43; Conservative 13; Mismatches 26; Indels 4; Gaps 1;

```

```

QY      1 LTLSPFOOQSLLDLVRAHTMAVNDGNETASDAGPSAFHIGFDSVWGVELRNL 60
          |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
          4762 LAALPEARGHLLLEVLAEFTASTLGH-----SAAVQPDRTFAELGFDSTLAVELRNL 4817
          :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY      61 SKATGLRLPVTLLFDHTTPAAVAARL 86
          :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB      4818 NAVTGLRLPVTLLFDHTPLALSEQ 4843

```

RESULT 5

```

US-09-105-537-6
; Sequence 6, Application US/09105537A
; Patent No. 6265202

```

```

; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6

```

```
/ LENGTH: 11877
/ TYPE: PRT
/ ORGANISM: Streptomyces venezuelae
US-09-105-537-6

Query Match
Best Local Similarity 46.9%; Score 199; DB 4; Length 11877;
Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;

QY 1 LTALSPTOOQSLLDLVRAHTMAVLNDGERTASDAGSPASPAHLGPDSPVNGVELRNRL 60
DB 2821 LAAAGGERTETILGLVRAQAAVL---RMRSPEDVADRAXKDIGPDLGVELRNRL 2876
QY 61 SKATGLRLPVTLIFDHTTPAAVAARL 86
DB 2877 TRATGLRLPATLVFDHPTPALVALSL 2902

RESULT 6
US-09-428-517-2
/ Sequence 2, Application US/09428517
/ Patent No. 6251636
/ GENERAL INFORMATION:
/ APPLICANT: Betlach, Mary C.
/ APPLICANT: Shah, Sanjay Krishnakant
/ APPLICANT: McDaniel, Robert
/ APPLICANT: Tang, Li
/ TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
/ FILE REFERENCE: 30062-20029.00
/ CURRENT APPLICATION NUMBER: US/09/428, 517
/ EARLIER FILING DATE: 1999-10-28
/ EARLIER APPLICATION NUMBER: 60/120,254
/ EARLIER FILING DATE: 1999-02-16
/ EARLIER APPLICATION NUMBER: 60/106,100
/ EARLIER FILING DATE: 1998-10-29
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 4150
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-428-517-2

Query Match
Best Local Similarity 46.2%; Score 196; DB 4; Length 4150;
Matches 43; Conservative 10; Mismatches 29; Indels 4; Gaps 1;

QY 1 LTALSPTOOQSLLDLVRAHTMAVLNDGERTASDAGSPASPAHLGPDSPVNGVELRNRL 60
DB 2441 LAGLTVAGGERLVLQVREGAAVVLGHSGEAVVA---PDRAKDXLGPDSLTVSELRNRL 2496
QY 61 SKATGLRLPVTLIFDHTTPAAVAARL 86
DB 2497 NTATGLRLPVTVAFDVARPALAHL 2522

RESULT 7
US-09-320-878-3
/ Sequence 3, Application US/09320878A
/ Patent No. 6117659
/ GENERAL INFORMATION:
/ APPLICANT: Ashley, Gary
/ APPLICANT: Betlach, Melanie C.
/ APPLICANT: McDaniel, Robert
/ APPLICANT: Tang, Li
/ TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
/ FILE REFERENCE: 300622002120
/ CURRENT APPLICATION NUMBER: US/09/320, 878A
/ CURRENT FILING DATE: 1999-05-27
```

```
/ EARLIER APPLICATION NUMBER: CIP OF 09/141,908
/ EARLIER FILING DATE: 1998-08-28
/ EARLIER APPLICATION NUMBER: CIP OF 09/073,538
/ EARLIER FILING DATE: 1998-05-06
/ EARLIER APPLICATION NUMBER: CIP OF 08/846,247
/ EARLIER FILING DATE: 1997-04-30
/ EARLIER APPLICATION NUMBER: 60/119,139
/ EARLIER FILING DATE: 1999-02-08
/ EARLIER APPLICATION NUMBER: 60/100,880
/ EARLIER FILING DATE: 1998-09-22
/ EARLIER APPLICATION NUMBER: 60/087,080
/ EARLIER FILING DATE: 1998-05-28
/ NUMBER OF SEQ ID NOS: 34
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 3
/ LENGTH: 1562
/ TYPE: PRT
/ ORGANISM: Streptomyces venezuelae
US-09-320-878-3

Query Match
Best Local Similarity 46.0%; Score 195; DB 3; Length 1562;
Matches 42; Conservative 11; Mismatches 29; Indels 4; Gaps 1;

QY 1 LTALSPTOOQSLLDLVRAHTMAVLNDGERTASDAGSPASPAHLGPDSPVNGVELRNRL 60
DB 1393 ITALPEPERPALTLVTRTHAAVVLGHSSPDRAVA---FGRAFTLGLPDSLTVAVQLRNOL 1448
QY 61 SKATGLRLPVTLIFDHTTPAAVAARL 86
DB 1449 STVGNRLPATVFDHPTPALAHL 1474

RESULT 8
US-09-105-537-35
/ Sequence 35, Application US/09105537A
/ Patent No. 6265202
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, D.H.
/ APPLICANT: Liu, H.
/ APPLICANT: Xue, Y.
/ APPLICANT: Zhao, L.
/ TITLE OF INVENTION: DNA encoding methymycin and pikromycin
/ FILE REFERENCE: 600.438US1
/ CURRENT APPLICATION NUMBER: US/09/105,537A
/ CURRENT FILING DATE: 1998-06-26
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 35
/ LENGTH: 1562
/ TYPE: PRT
/ ORGANISM: Streptomyces venezuelae
US-09-105-537-35

Query Match
Best Local Similarity 46.0%; Score 195; DB 4; Length 1562;
Matches 42; Conservative 11; Mismatches 29; Indels 4; Gaps 1;

QY 1 LTALSPTOOQSLLDLVRAHTMAVLNDGERTASDAGSPASPAHLGPDSPVNGVELRNRL 60
DB 1393 ITALPEPERPALTLVTRTHAAVVLGHSSPDRAVA---FGRAFTLGLPDSLTVAVQLRNOL 1448
QY 61 SKATGLRLPVTLIFDHTTPAAVAARL 86
DB 1449 STVGNRLPATVFDHPTPALAHL 1474

RESULT 9
US-09-036-987A-3
/ Sequence 3, Application US/09036987A
/ Patent No. 6143526
/ GENERAL INFORMATION:
/ APPLICANT: Baltz, Richard H.
```

```

/ APPLICANT: Broughton, Mary C.
/ APPLICANT: Crawford, Kathryn P.
/ APPLICANT: Madduri, Krishnamurthy
/ APPLICANT: Merlo, Donald J.
/ APPLICANT: Tredey, Pat J.
/ APPLICANT: Turner, Jan R.
/ APPLICANT: Waldron, Clive
/ TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
/ TITLE OF INVENTION: Production
/ NUMBER OF SEQUENCES: 39
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Dow Agrosiences LLC Patent Department
/ STREET: 9330 Zionsville Road
/ CITY: Indianapolis
/ STATE: Indiana
/ COUNTRY: USA
/ ZIP: 46268
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/036,987A
/ FILING DATE: 09-MAR-1998
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Stuart, Donald R.
/ REGISTRATION NUMBER: 28,479
/ REFERENCE/DOCKET NUMBER: 50,608
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (317)337-4816
/ TELEFAX: (317)337-4847
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2152 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-036-987A-3

Query Match          45.3%; Score 192; DB 4; Length 2152;
Best Local Similarity 50.0%; Pred. No. 2.3e-16;
Matches 43; Conservative 13; Mismatches 26; Indels 4; Gaps 1;

QY 1 LTALEPTQOQSLLDLVRAHTMAVINDGNERITASDAGPSAFHIGFDSVNGVELRNRL 60
DB 1989 LGRLPBAEQHRLLDLVRHVAALVGFASQSEITD---GTFKVLGFSLTVVELRNRI 2044

QY 61 SKATGRLPVTLIFDHTTPAAVAARL 86
DB 2045 NGATGRLPATLVFNPTPDALAHL 2070

RESULT 10
US-09-370-700-3
/ Sequence 3, Application US/09370700
/ Patent No. 6274350
/ GENERAL INFORMATION:
/ APPLICANT: Baltz, Richard H.
/ APPLICANT: Broughton, Mary C.
/ APPLICANT: Crawford, Kathryn P.
/ APPLICANT: Madduri, Krishnamurthy
/ APPLICANT: Tredey, Pat J.
/ APPLICANT: Turner, Jan R.
/ APPLICANT: Waldron, Clive
/ TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
/ FILE REFERENCE: 50489 DIV1
/ CURRENT APPLICATION NUMBER: US/09/370,700
/ CURRENT FILING DATE: 1999-08-09
/ EARLIER APPLICATION NUMBER: US 09/36987
/ EARLIER FILING DATE: 1998-03-09
/ NUMBER OF SEQ ID NOS: 39
```

```

/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 3
/ LENGTH: 2152
/ TYPE: PRT
/ ORGANISM: Saccharopolyspora spinosa
/ US-09-370-700-3

Query Match          45.3%; Score 192; DB 4; Length 2152;
Best Local Similarity 50.0%; Pred. No. 2.3e-16;
Matches 43; Conservative 13; Mismatches 26; Indels 4; Gaps 1;

QY 1 LTALEPTQOQSLLDLVRAHTMAVINDGNERITASDAGPSAFHIGFDSVNGVELRNRL 60
DB 1989 LGRLPBAEQHRLLDLVRHVAALVGFASQSEITD---GTFKVLGFSLTVVELRNRI 2044

QY 61 SKATGRLPVTLIFDHTTPAAVAARL 86
DB 2045 NGATGRLPATLVFNPTPDALAHL 2070

RESULT 11
US-08-804-227C-10
/ Sequence 10, Application US/08804227C
/ Patent No. 5876991
/ GENERAL INFORMATION:
/ APPLICANT: Dehoff, Bradley S.
/ APPLICANT: Kuhstoss, Stuart A.
/ APPLICANT: Rostock, Paul R., Jr.
/ APPLICANT: Sutton, Kimberly L.
/ TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
/ NUMBER OF SEQUENCES: 15
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: THOMAS G. PLANT 1501
/ STREET: LILLY CORPORATE CENTER
/ CITY: INDIANAPOLIS
/ STATE: IN
/ COUNTRY: USA
/ ZIP: 46285
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: ASCII (DOS) Text only
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/804,227C
/ FILING DATE: February 21, 1997
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Plant, Thomas G.
/ REGISTRATION NUMBER: 35,784
/ REFERENCE/DOCKET NUMBER: X-8231
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 317-276-2459
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3724 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: Peptide
/ US-08-804-227C-10

Query Match          44.8%; Score 190; DB 2; Length 3724;
Best Local Similarity 47.7%; Pred. No. 9.2e-16;
Matches 41; Conservative 17; Mismatches 24; Indels 4; Gaps 1;

QY 1 LTALEPTQOQSLLDLVRAHTMAVINDGNERITASDAGPSAFHIGFDSVNGVELRNRL 60
DB 1440 LTAQGHKRAHTLTLVRAHAAALVGGSSGDAVSS---ARAFRLGFSLTVLALRLRL 1495

QY 61 SKATGRLPVTLIFDHTTPAAVAARL 86
DB 1496 STGTGKLPTSLVFNHSSPAALARL 1521
```


RESULT 12
US-08-804-198-4
Sequence 4, Application US/08804198
Patent No. 5945320
GENERAL INFORMATION:
APPLICANT: Burgett, Stanley G.
APPLICANT: Kuhse, Stuart A.
APPLICANT: Rao, Nagaraja R.
APPLICANT: Richardson, Mark A.
APPLICANT: Rostek, Paul R., Jr.
TITLE OF INVENTION: PLATENOIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-4

Query Match 44.8%; Score 190; DB 2; Length 3724;
Best Local Similarity 47.7%; Pred. No. 9.2e-16;
Matches 41; Conservative 17; Mismatches 24; Indels 4; Gaps 1;

QY 1 LTAISPTQOQSLLDLVRAHTMAVNDGNERATSDAGPSAPAHGFDPSVWGVLELRNL 60
DB 1440 LTAOSGHKRAVTLRLVRAHAAVVGQSGDAVSS---ARAFRDLGFDLTLALRL 1495

QY 61 SKATGLRPLVTLLFDHTTPAAVAAARL 86
DB 1496 STGTGLKPLTSLVFDHSSPALAARL 1521

RESULT 13
US-08-804-227C-4
Sequence 4, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhse, Stuart A.
APPLICANT: Rostek, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA

ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3729 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-804-227C-4

Query Match 44.8%; Score 190; DB 2; Length 3729;
Best Local Similarity 47.7%; Pred. No. 9.2e-16;
Matches 41; Conservative 12; Mismatches 29; Indels 4; Gaps 1;

QY 1 LTAISPTQOQSLLDLVRAHTMAVNDGNERATSDAGPSAPAHGFDPSVWGVLELRNL 60
DB 3547 LTGLTAOBGHALVLEIRAHAAVVGHGSDSIPED---RAFKDGLGFDLTLAVENRNL 3602

QY 61 SKATGLRPLVTLLFDHTTPAAVAAARL 86
DB 3603 SAATGLRPLATVTFDHPPELGH 3628

RESULT 14
US-09-105-537-2
Sequence 2, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438US1
CURRENT APPLICATION NUMBER: US/09/105,537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PaeSeQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 5215
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-2

Query Match 43.2%; Score 183; DB 4; Length 5215;
Best Local Similarity 50.0%; Pred. No. 1.3e-14;
Matches 43; Conservative 13; Mismatches 26; Indels 4; Gaps 2;

QY 1 LTAISPTQOQSLLDLVRAHTMAVNDGNERATSDAGPSAPAHGFDPSVWGVLELRNL 60
DB 2713 LAAQSPERRSTAVTLVTVGVADV--GHAUSA-VGAERFQKDLGFDLTLAVENRNL 2768

QY 61 SKATGLRPLVTLLFDHTTPAAVAAARL 86
DB 2769 NATGLRPLATVTFDHPSPALASHL 2794

RESULT 15
US-09-320-878-2

```
; Sequence 2, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: McDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 3739
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-320-878-2
```

```
Query Match          42.9%; Score 182; DB 3; Length 3739;
Best Local Similarity 47.7%; Pred. No. 1,1e-14;
Matches 41; Conservative 9; Mismatches 32; Indels 4; Gaps 1;

QY      1 LTLSPTOQOSLLDLVRAHTMAVLNDGNERITASDAGPSASFAHLGFDSEVMGVELRNRL 60
         : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      3560 LAAMTPDDRVAHRLDLVRTVAATVL---GHGTPSRVDLERAFRDGTGFDSTLTAVELRNRL 3615
         : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      61 SKATGRLPYTLIEDHTPAVAARL 86
         : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      3616 NAATGRLPYTLVFDHTPGELAGHL 3641
         : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: June 17, 2003, 13:09:25
Job time : 2.5631 secs

SPD a polyketide
S. venezuelae pik
S. antibioticus 8
Nardoniolide synth
Amino acid sequen
S. venezuelae maci
S. venezuelae pik
Amycolatopsis med
A. mediterranei r
S. avermitilis av
Streptomyces aver
SpnB a polyketide
SpnB a polyketide
S. spinosa proteol
Platanolide synth
Platanolide synth
Tylosone synthase
Streptomyces nour
A. mediterranei r
A. mediterranei r
A. mediterranei r
S. venezuelae vep
Amino acid sequen
S. venezuelae maci
S. venezuelae pik
Nardoniolide synth
SpnC a polyketide
S. spinosa proteol
Polyketide synthas
Streptomyces nour
Mycobacterium tub
Pimaricin biosynt
Tylosone synthase
Polyketide synthas
Streptomyces vena
S. avermitilis av

PT agrochemical use -
 XX
 PS Claim 32; Page 203-223; 314pp; Japanese.
 XX
 CC The present sequence represents an avermectin aglycon synthase protein.
 CC Also described are: (1) polypeptides encoded by all or part of the DNA;
 CC (2) expression vectors containing the DNA; (3) host cells transformed by
 CC the vectors; (4) preparation of the polypeptides by culture of the
 CC transformants; (5) preparation of avermectin aglycon or its derivatives
 CC by culture of transformed avermectin-producing microorganisms; and (6)
 CC oligonucleotides of 5-60 bases in length containing sense or antisense
 CC sequences from the avermectin aglycon synthase DNA. The enzymes are
 CC useful for the production of modified forms of avermectin and of the
 CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
 CC and agrochemicals.
 CC
 SQ Sequence 3972 AA;
 Query Match 100.0%; Score 424; DB 21; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 4.6e-46;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LTLSPQQOQSLLDLVRAHTMAVLNDGNETASDAGPSAFHIGFDSVWGVELRNRL 60
 DB 366 LTLSPQQOQSLLDLVRAHTMAVLNDGNETASDAGPSAFHIGFDSVWGVELRNRL 425
 QY 61 SKATGLRPVTLIFDHTTPAAVAARL 86
 DB 426 SKATGLRPVTLIFDHTTPAAVAARL 451
 RESULT 2
 AAG65264
 ID AAG65264 standard; Protein; 3972 AA.
 AC AAG65264;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Streptomyces avermectilis protein SEQ ID NO: 4.
 XX
 KW Avermectin aglycone synthase; AAS; avermectin derivative;
 KW drug production; veterinary drug; pesticide.
 XX
 OS Streptomyces avermectilis.
 XX
 PN MO200162939-A1.
 PD 30-AUG-2001.
 PF 23-FEB-2001; 2001MO-JP01381.
 PR 24-FEB-2000; 2000JP-0047405.
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 PA (KITA) KITASATO INST.
 PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
 DR WPI; 2001-582053/65.
 DR N-PSDB; AAH79277.
 PT New modified avermectin aglycone synthase derived from Streptomyces
 PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
 PT drugs and pesticides -
 CC
 XX Claim 4; Page 167-180; 257pp; Japanese.
 CC The present invention relates to the production of modified derivatives
 CC of avermectin aglycone synthase (AAS) derived from Streptomyces
 CC avermectilis. The activity of an acyl carrier protein (ACP),
 CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
 CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl

CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is an S. avermectilis protein.
 CC
 SQ Sequence 3972 AA;
 Query Match 100.0%; Score 424; DB 22; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 4.6e-46;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LTLSPQQOQSLLDLVRAHTMAVLNDGNETASDAGPSAFHIGFDSVWGVELRNRL 60
 DB 366 LTLSPQQOQSLLDLVRAHTMAVLNDGNETASDAGPSAFHIGFDSVWGVELRNRL 425
 QY 61 SKATGLRPVTLIFDHTTPAAVAARL 86
 DB 426 SKATGLRPVTLIFDHTTPAAVAARL 451
 RESULT 3
 AAG65268
 ID AAG65268 standard; Protein; 3972 AA.
 AC AAG65268;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Streptomyces avermectilis protein derivative SEQ ID NO: 8.
 XX
 KW Avermectin aglycone synthase; AAS; avermectin derivative;
 KW drug production; veterinary drug; pesticide.
 XX
 OS Synthetic.
 XX
 PN MO200162939-A1.
 PD 30-AUG-2001.
 PF 23-FEB-2001; 2001MO-JP01381.
 PR 24-FEB-2000; 2000JP-0047405.
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 PA (KITA) KITASATO INST.
 PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
 DR WPI; 2001-582053/65.
 DR N-PSDB; AAH79277.
 PT New modified avermectin aglycone synthase derived from Streptomyces
 PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
 PT drugs and pesticides -
 CC
 XX Claim 5; Page 235-248; 257pp; Japanese.
 CC The present invention relates to the production of modified derivatives
 CC of avermectin aglycone synthase (AAS) derived from Streptomyces
 CC avermectilis. The activity of an acyl carrier protein (ACP),
 CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
 CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is derived from an S.
 CC avermectilis protein.
 CC
 SQ Sequence 3972 AA;
 Query Match 100.0%; Score 424; DB 22; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 4.6e-46;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LTLSPQQOQSLLDLVRAHTMAVLNDGNETASDAGPSAFHIGFDSVWGVELRNRL 60

DB 366 LTRALSPFOOQSLLDLVRAHTMAVLNDGNETASDAGSPASPAHLGFPDSVMGVELRNRL 425

QY 61 SKATGLRLPVTLLFDHTTPAAVAARL 86
DB 426 SKATGLRLPVTLLFDHTTPAAVAARL 451

RESULT 4

ID AAE10145 standard; Protein; 5435 AA.

AC AAE10145;

DT 29-NOV-2001 (first entry)

DE Streptomyces noursei nystatin gene, NysJ.

KW Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;

KM antifungal; antibiotic.

OS Streptomyces noursei.

PN WO200159126-A2.

PD 16-AUG-2001.

PF 08-FEB-2001; 2001WO-GB00509.

PR 08-FEB-2000; 2000GB-0002840.

PR 10-APR-2000; 2000GB-0008786.

PR 14-APR-2000; 2000GB-0009387.

PA (UNO-) UNIV NORGE TEKNIKSK NATURVITENSKABELIGE.

PA (SNP) SINTEF STIFTELSEN IND TEK FORSK.

PA (ALP-) ALPHARMA AS.

PA (SINV-) SINVENT AS.

PA (DZIE/) DZIELEMSKA H.

PA (ZOTC/) ZOTCHEV S B.

PA (SEKU/) SEKUROVA O N.

PA (FJAE/) FJAEVRIK E.

PA (BRAU/) BRAUTASET T.

PA (STRO/) STROM A R.

PI Zolchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;

PI Valla S, Ellingsen TE, Sletta H, Gulliksen O;

DR MPI; 2001-557614/62.

DR N-PSDB; AAD17186.

XX New nystatin polyketide synthase polynucleotides and polypeptides.

PT useful as antibiotics and antifungals -

PS Claim 15; Page 260-263; 266pp; English.

XX The present invention relates to the cloning and sequencing of the gene

CC cluster encoding a modular type I polyketide synthase (PKS) enzyme

CC involved in the biosynthesis of the macrolide antibiotic nystatin.

CC The nystatin PKS is useful as antifungal antibiotic. The present

CC sequence is a PKS encoding Streptomyces noursei nystatin gene, NysJ.

DB 5320 LATATGARLPATLVFDHTTPAAVAARL 5346

RESULT 5

ID AAM47200 standard; Peptide; 47 AA.

AC AAM47200;

DT 12-FEB-2002 (first entry)

DE Modular enzyme system related ACP-domain N-terminal peptide AVEA1_1.

KW Modular enzyme system; cyclic gene synthesis; repetitive coding sequence;

KM antibiotic; non-ribosomal peptide synthetase; NRPS; PKS;

OS Streptomyces avermitilis.

PN WO200181564-A2.

PD 01-NOV-2001.

PF 25-APR-2001; 2001WO-DE01578.

PR 26-APR-2000; 2000DE-1021267.

PA (ACTI-) ACTINODRUG PHARM GMBH.

PI Schaecker F;

DR MPI; 2002-049276/06.

PT Preparing DNA encoding modular protein for e.g. producing new enzymes

PT for synthesis of polyketide antibiotics, comprises cyclic integration

PT of fragments into a vector -

XX Example 1; Fig 9; 83pp; German.

CC The present invention relates to the preparation of DNA, in a circular

CC vector, that encodes one or more segments of a modular polypeptide. DNA

CC or DNA libraries produced this way are used to produce modular

CC polypeptides, particularly enzymes, which can be used to act on

CC substrates to produce compounds for therapeutic testing. Enzymes of

CC particular interest are those involved in non-ribosomal peptide synthesis

CC or polyketide synthesis, and compounds for testing are particularly

CC macrolide antibiotics, including penicillins, vancomycin or

CC erythromycin, but may also be modular receptors. The present sequence is

XX a peptide used in the exemplification of the invention.

SO Sequence 47 AA;

Query Match 47.1%; Score 199.5; DB 23; Length 47;

Best Local Similarity 89.6%; Pred. No. 66-19; Indels 5; Gaps 1;

Matches 43; Conservative 0; Mismatches 0;

QY 39 PSASPAHLGFPDSVMGVELRNRLSKATGLRLPVTLLFDHTTPAAVAARL 86

DB 1 PSASPAHLGFPDSVMGVELRNRLSKATGLRLPVTLLFDHTTPAAVAARL 43

RESULT 6

ID AAB18637 standard; Protein; 4551 AA.

AC AAB18637;

DT 22-JAN-2001 (first entry)

DE Amino acid sequence of narbonolide synthase subunit 1 (PICA1).

KW Narbonolide synthase; polyketide synthase gene; narbonolide polyketide;

KM antibiotic; C12-hydroxylase; pick; deosamine biosynthesis;

KM desosaminyl transferase enzyme; ketolide; beta-glucosidase enzyme;
 XX picromycin biosynthesis.
 XX OS Streptomyces venezuelae.
 XX PN US6117659-A.
 XX PD 12-SEP-2000.
 XX PF 27-MAY-1999; 99US-0320878.
 XX PR 28-MAY-1998; 98US-0087080.
 XX PR 22-SEP-1998; 98US-0100880.
 XX PR 08-FEB-1999; 99US-0119139.
 XX PR 20-MAY-1999; 99US-0134990.
 XX PR 30-APR-1997; 97US-0846247.
 XX PR 06-MAY-1998; 98US-0073538.
 XX PR 28-AUG-1998; 98US-0141908.
 XX PA (KOSA-) KOSAN BIOSCIENCES INC.
 XX PI Ashley G, Betlach MC, Betlach M, Tang L, McDaniel R;
 XX DR WPI; 2000-610844/58.
 XX PT New recombinant pick hydroxylase gene of Streptomyces venezuelae useful
 XX PT for converting ketolides to antibiotics and as antibiotics and
 XX PT intermediates in the synthesis of compounds with pharmaceutical value
 XX PT -
 XX PS Disclosure; Columns 9-10; 117pp; English.
 XX SX
 CC The present sequence represents a narbonolide synthase subunit 1
 CC (PICAI). The nucleotide sequence encoding it is used in the course of
 CC the invention. The specification describes a recombinant DNA compound
 CC expressing recombinant polyketide synthase genes in host cells for the
 CC production of narbonolide, narbonolide derivatives and polyketides that
 CC are useful as antibiotics and as intermediates in the synthesis of
 CC compounds with pharmaceutical value. The DNA compounds may also encode
 CC a C12-hydroxylase (pick), desosamine biosynthesis and desosaminyl
 CC transferase enzymes (useful for conversion of ketolides to antibiotics),
 CC and the beta-glucosidase enzyme (involved in picromycin biosynthesis).
 CC These compounds are also useful for increasing the antibiotic activity
 CC of a compound relative to the unhydroxylated compound. The recombinant
 CC host cells are useful as genetic systems that allow rapid engineering
 CC of the narbonolide polyketide synthase. These would be valuable for
 CC creating novel ketolide analogs for pharmaceutical applications.
 CC CX
 CC SX
 SQ Sequence 4551 AA;
 Query Match 46.9%; Score 199; DB 21; Length 4551;
 Best Local Similarity 48.8%; Pred. No. 3.5e-16;
 Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;
 QY 1 LTAASPTQOOSLLDLVRAHTVAVLNDGNERFASDAGSASFALHGFDSVNGVELRNRL 60
 DB 2423 LAAAPGERTREILGLVRAQAALV----RMSPEVADRAKFDIGFSLAGVELRNRL 2478
 QY 61 SKATGRLPVTLLFDHTTPAANAAL 86
 DB 2479 TRATGQLPATLVFDHPTPLALVSL 2504
 RESULT 7
 ID AAY67201 standard; protein; 4551 AA.
 AC AAY67201;
 XX
 XX
 DT 23-MAR-2000 (first entry)
 XX
 DE Narbonolide synthase subunit 1 (PICAI) protein sequence.
 XX

KM Narbonolide polyketide synthase; PKS; narbonolide synthase subunit 1;
 KM PICAI; antibiotic production; narbomycin; picromycin; ketolide.
 XX OS Streptomyces venezuelae.
 XX PN MO961599-A2.
 XX PD 02-DEC-1999.
 XX PF 27-MAY-1999; 99WO-US11814.
 XX PR 28-MAY-1998; 98US-0087080.
 XX PR 28-AUG-1998; 98US-0141908.
 XX PR 22-SEP-1998; 98US-0100880.
 XX PR 08-FEB-1999; 99US-0119139.
 XX PA (KOSA-) KOSAN BIOSCIENCES INC.
 XX PI Ashley G, Betlach MC, Betlach M, McDaniel R, Tang L;
 XX DR WPI; 2000-072618/06.
 XX DR N-PSDB; AAZ56001.
 XX PT New recombinant DNA encoding a domain of narbonolide polyketide
 XX PT synthase, for production of ketolide antibiotics -
 XX PS Example 2; Page 11-12; 98pp; English.
 XX SX
 CC This is the Streptomyces venezuelae narbonolide synthase subunit 1,
 CC PICAI protein sequence. The invention relates to recombinant DNA
 CC containing a coding sequence for a narbonolide polyketide synthase
 CC (PKS). Polyketides are compounds synthesised from 2-carbon units through
 CC a series of condensations and subsequent modifications. Modular PKSs are
 CC responsible for the production of many antibiotics including picromycin.
 CC The narbonolide PKS consists of a loading module, six extender modules,
 CC and two thioester domains. Four proteins make up the narbonolide PKS
 CC (PICAI, PICAILI, PICALII and PICALIV). PICAI includes the loading module
 CC and extender modules 1 and 2, PICALII includes extender modules 3 and 4,
 CC PICALIII includes extender module 5 and PICALIV includes extender module 6
 CC and a type II thioesterase domain. The second type II thioesterase
 CC domain is found on the PICB protein. The nucleotide sequences encoding
 CC all of these proteins can be isolated in recombinant form from the
 CC recombinant cosmid pKOS023-27 (see AAZ56001). Narbonolide is
 CC desosaminylated in S. venezuelae to yield narbomycin, the desosaminyl
 CC transferase enzyme is required for this conversion, and the recombinant
 CC biosynthetic genes are also found in cosmid pKOS023-27. The recombinant
 CC DNA of the invention is used to express, in transformed cells,
 CC narbonolide (or its derivatives) or other ketolides (particularly
 CC hybrids), which may then be converted (e.g. by other enzymes
 CC recombinantly expressed in the same hosts) to polyketide antibiotics or
 CC their intermediates. The antibiotics are useful in human or veterinary
 CC medicine.
 CC CX
 CC SX
 SQ Sequence 4551 AA;
 Query Match 46.9%; Score 199; DB 21; Length 4551;
 Best Local Similarity 48.8%; Pred. No. 3.5e-16;
 Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;
 QY 1 LTAASPTQOOSLLDLVRAHTVAVLNDGNERFASDAGSASFALHGFDSVNGVELRNRL 60
 DB 2423 LAAAPGERTREILGLVRAQAALV----RMSPEVADRAKFDIGFSLAGVELRNRL 2478
 QY 61 SKATGRLPVTLLFDHTTPAANAAL 86
 DB 2479 TRATGQLPATLVFDHPTPLALVSL 2504
 RESULT 8
 ID AAY77192 standard; Protein; 4613 AA.
 AC AAY77192;
 XX
 XX

[illegible]

QY 1 LTALSPFOQSLDDLVLRATMTAVLNDGDIERFASDPGPEASPAHLQFDSVMGVELRNRL 60
 Db 2485 LAAAPAEERREIILLGLVRAQAANVL-----RMRSPEDYAAADRAKRDICFDSLAGVELRNRL 2540
 QY 61 SKATGLRLPVTLIFDHTTPAAVAARL 86
 Db 2541 TRATGLQLPATLVFDHPTPLALVSL 2566
 RESULT 9
 AAY77200
 XX AAY77200 standard; Protein; 4613 AA.
 AC AAY77200;
 XX AAY77200;
 DT 05-JUN-2000 (first entry)
 XX
 DE 5. venezuelae pik (macrolide biosynthesis) gene cluster protein #1.
 XX
 KW Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
 KW neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
 KW biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;
 KW chronic obstructive pulmonary disease; respiratory inflammation;
 KW hypercholesterolaemia; crop protection agent.
 XX
 OS Streptomyces venezuelae ATCC15439.
 XX
 PN W0200000620-A2.
 XX
 PD 06-JAN-2000.
 XX
 PF 25-JUN-1999; 99WO-US14398.
 XX
 PR 26-JUN-1998; 98US-0105537.
 XX
 XX (MINU) UNIV MINNESOTA.
 PA
 PI Sherman DH, Liu H, Xue Y, Zhao L;
 DR WPI; 2000-160679/14.
 DR N-PSDS; AAZ87318.
 PT
 PT Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.
 PS synthesis of methymycin and pikromycin -
 XX
 XX Disclosure; Figure 31; 438bp; English.
 CC The invention relates to an isolated and purified nucleic acid segment
 CC comprising a desosamine biosynthetic gene cluster, a fragment or its
 CC biologically active variant, where the nucleic acid sequence is not
 CC derived from the eycC gene cluster of *Saccharopolyspora erythraea* or
 CC streptomycetes antibiotics. The invention also relates to a macrolide
 CC biosynthetic gene cluster, or fragments thereof. The macrolide
 CC biosynthetic gene cluster encodes proteins which synthesise methymycin,
 CC pikromycin, neomethymycin, narbomycin or a combination of these
 CC compounds. Recombinant or augmented cells comprising the desosamine
 CC and/or macrolide biosynthetic gene clusters are useful for the
 CC production of biologically active macrolides. The macrolide biosynthetic
 CC proteins are useful for synthesis of methymycin, pikromycin,
 CC neomethymycin and narbomycin. The alternative termination of polyketide
 CC synthesis may be useful to prepare novel antibiotics and
 CC polyhydroxyalkanoate (PHA) monomers. The compounds produced by the
 CC recombinant host cells are useful as biopolymers, e.g., in packaging or
 CC biomedical applications, to engineer PHA monomer synthases or to prepare
 CC biologically active agents, such as chemotherapeutics,
 CC immunosuppressants, agents to treat asthma, chronic obstructive pulmonary
 CC disease as well as other diseases involving respiratory inflammation,
 CC cholesterol-lowering agents or macrolide-based antibiotics which are
 CC active against a variety of organisms, e.g., bacteria, including
 CC multi-drug resistant pneumococci and other respiratory pathogens, as well
 CC as viral parasitic pathogens, or as crop protection agents (e.g.,
 CC fungicides or insecticides) via expression of polyketides in plants.

CC Sequences AAY77200-Y77203 and AAY80997 represent proteins encoded by the
 CC macrolide biosynthetic gene cluster (pik) from Streptomyces venezuelae
 CC ATCC 15439, as given in figure 31.
 XX
 SQ Sequence 4613 AA;
 Query Match 46.9%; Score 199; DB 21; Length 4613;
 Best Local Similarity 48.8%; Pred. No. 3.6e-16;
 Matches 42; Conservative 14; Mismatches 26; Indels 4; Gaps 1;
 QY 1 LTLSPFOOQSLDLVRAHTMAVINDGNERASDAGPSAFHIGFDSVNGVELRNRL 60
 DB 2485 LAAAPRERTELLGLVRAQAAVLT---RMRSPEDVADRAKDIGFDSLAVGVELRNRL 2540
 QY 61 SKATGRLPVTLLIFDHTTPAAVAA 86
 DB 2541 TRATGQLPVTLLVFDHTTPALVSL 2566
 RESULT 10
 AAB70968
 ID AAB70968 standard; Protein; 4924 AA.
 AC AAB70968;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 DE S. spinosa protein fragment encoded by ORF21, SEQ ID 48.
 KW Forosamine, trimethylhamose; polyketide synthase; biosynthesis;
 KW spinosyn; polyketide aglycone; transgenic plant; insect resistance;
 KW macrolide; insecticidal; polyketide synthase.
 XX
 OS Saccharopolyspora spinosa.
 PN DE19957268-A1.
 XX
 PD 08-MAR-2001.
 XX
 PF 29-NOV-1999; 99DE-1057268.
 XX
 PR 27-AUG-1999; 99DE-1040596.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Eberz G, Moehle V, Froede R, Velten R, Salas UA;
 XX
 DR WPI, 2001-267102/28.
 DR N-PSDB; AAF88338.
 XX
 PT New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
 PT recombinant production of insecticidal spinosyns and their derivatives
 XX
 PS Claim 57; Page 264-284; 354pp; German.
 XX
 CC This invention describes a novel method nucleic acid (I) and its encoded
 CC polypeptide (II) containing at least one region that encodes an enzymatic
 CC activity involved in biosynthesis of spinosyns. (I) are used (i) to
 CC identify, inactivate or modulate genes involved in the biosynthesis of
 CC (II); (ii) to generate a library of polyketide synthases; (iii) for
 CC adding forosamine or trimethylhamose to a spinosyn or polyketide
 CC aglycone; and (iv) for recombinant production of the corresponding
 CC enzymes, which are used for production of (II), their precursors or
 CC derivatives, including production of transgenic plants that express (II)
 CC and thus have increased resistance to insects. (I) are also useful as
 CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
 CC macrolides with insecticidal, but not antibacterial, activity, and can
 CC also be used to raise specific antibodies, useful for identifying
 CC expression clones in a gene bank. Cells transformed with (I) may produce
 CC (II) at significantly increased levels or produce new derivatives of
 CC (II). This sequence represents an S. spinosa polyketide synthase.
 XX

SQ Sequence 4924 AA;
 Query Match 46.9%; Score 199; DB 22; Length 4924;
 Best Local Similarity 50.0%; Pred. No. 3.9e-16;
 Matches 43; Conservative 13; Mismatches 26; Indels 4; Gaps 1;
 QY 1 LTLSPFOOQSLDLVRAHTMAVINDGNERASDAGPSAFHIGFDSVNGVELRNRL 60
 DB 4758 LAAAPRERGHILLEVLAEFTASTLGHGHD---SAAVQPDRTFAELGFDSLAVGVELRNRL 4813
 QY 61 SKATGRLPVTLLIFDHTTPAAVAA 86
 DB 4814 NAVTGRLEPPTLVFDPHPTLALSEQL 4839
 RESULT 11
 AAY39300
 ID AAY39300 standard; Protein; 4928 AA.
 AC AAY39300;
 XX
 DT 01-DEC-1999 (first entry)
 XX
 DE Spnd a polyketide synthase.
 XX
 KW Spinosyn biosynthetic enzyme; open reading frame; ORF; insecticidal
 KW macrolides; arachnid; nematode; insect; polyketide; polyketide synthase;
 KW PKS; extender module; initiator module; acyl transferase domain; AT;
 KW acyl carrier protein; ACP; beta-ketosynthase domain; KS; KR;
 KW dehydratase domain; DH; enoyl reductase domain; ER; beta-ketoreductase;
 KW insecticide.
 XX
 OS Saccharopolyspora spinosa.
 PN
 XX
 FH Key
 FT Domain
 FT Location/Qualifiers
 FT 1..424
 FT /label= KS5
 FT /note= "Beta-ketosynthase domain: part of extender
 FT module 5"
 FT 539..866
 FT /label= AT5
 FT /note= "Acyl transferase domain: part of extender
 FT module 5"
 FT 893..1078
 FT /label= DH5
 FT /note= "Dehydratase domain: part of extender
 FT module 5"
 FT 1384..1565
 FT /label= KR5
 FT /note= "Beta-ketoreductase domain: part of extender
 FT module 5"
 FT 1645..1726
 FT /label= ACP5
 FT /note= "Acyl carrier protein domain: part of extender
 FT module 5"
 FT 1748..2172
 FT /label= KS6
 FT /note= "Beta-ketosynthase domain: part of extender
 FT module 6"
 FT 2283..2613
 FT /label= AT6
 FT /note= "Acyl transferase domain: part of extender
 FT module 6"
 FT 2916..3095
 FT /label= KR6
 FT /note= "Beta-ketoreductase domain: part of extender
 FT module 6"
 FT 3188..3269
 FT /label= ACP6
 FT /note= "Acyl carrier protein domain: part of extender
 FT module 6"
 FT 3291..3713
 FT /label= KS7
 FT Domain

FT	/notes= "beta-ketosynthase domain: part of extender module 7"
FT	
FT	Domain
FT	3825..4153
FT	/label= AT7
FT	/note= "acyl transferase domain: part of extender module 7"
FT	
FT	Domain
FT	4344..4638
FT	/label= KR7
FT	/note= "beta-ketoreductase domain: part of extender module 7"
FT	
FT	Domain
FT	4725..4806
FT	/label= ACP7
FT	/note= "acyl carrier protein domain: part of extender module 7"
FT	
PX	M09946387-A1.
PX	
PD	16-SEP-1999.
PF	16-FEB-1999; 99WC-US03212.
PX	
PX	09-MAR-1998; 98US-0036987.
PR	(DOWC) DOW AGROSCIENCES LLC.
PA	
P1	Baltz RH, Broughton MC, Crawford KP, Madhuri K, Merlo DJ,
P1	Treadway PJ, Turner JR, Waldron C;
XK	WPI, 1999-551414/46.
DR	N-PSDB; AAZ21501.
XX	
PT	New spinosyn biosynthetic genes from Saccharopolyspora spinosa, useful for production of insecticidal spinosyn compounds -
PS	Claim 1; Page 113-129; 190pp; English.
XX	
CC	This is the amino acid sequence of the product of the spnd gene. The protein is involved in spinosyn biosynthesis. The spnd gene is one of 23 genes and open reading frames contained in an 80kb DNA sequence AAZ21501. Spinosyns are insecticidal macroides which are useful for the control of arachnids, nematodes and insects. Biosynthesis of spinosyns occurs via stepwise condensation and modification of carboxylic acid precursors generating a linear polyketide which is modified further. The DNA sequence contains a central region of approximately 55kb which has homology to the DNA encoding the polyketide synthases (PKS) of known macrolide producers. The spinosyn PKS DNA region consists of 5 ORFs with stop codons at the end of acyl carrier protein (ACP) domain. Together the PKS polypeptides (AAVY9297-Y39301) of which this sequence is one, form a complex consisting of an initiator module, spna, and several extender modules spnb-spm. Each extender module adds a specific acetyl-Co-A precursor to a growing polyketide chain, and modifies the beta-keto group in a specific manner. A module in a PKS polypeptide consists of several domains with specific functions. The initiator module has an acyl transferase (AT) domain, and an acyl carrier protein (ACP) domain. The extender modules have the same domain plus a beta-ketosynthase (KS) domain and optionally a beta-ketoreductase domain, a dehydratase (DH) domain, and an enoyl reductase (ER) domain. The last extender module terminates with a thioester domain. The products of the genes present in the upstream region the PKS genes have been assigned names spnr-spns AAVY9302-Y39315 and are responsible for different modifications in spinosyn biosynthesis. There are also two ORFs ORP15 and ORP16 present immediately upstream of spns, producing polypeptides AAVY9316-Y39317, and two ORFs ORP1 and ORP2 present downstream of the PKS region producing polypeptides AAVY9318-Y39319. The genes are useful to improve yields of spinosyns, and for creating new spinosyns e.g. by mutagenesis, or interruption of steps in spinosyn biosynthesis. The modified spinosyns may be a new insect control agent or serve as substrates for further chemical modification and the creation of new semi-synthetic spinosyns. The genes are also useful to isolate similar sequences from S spinosa or other species by hybridization.
Sequence	4928 AA;

```

Query Match          46.94; Score 199; DB 20; Length 4928;
Best Local Similarity 50.04; Pred. No. 3.9e-16;
Matches 43; Conservative 13; Mismatches 26; Indels 4; Gaps 1;

QY.      1 LTLSPFOOSSLIDLVRATMAVLNDGNGERTASDAGPSAFHJLFGDSVMGYENRL 60
      ||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      4762 LALLPRAERGHLLLEVLALETASTLGHG---SBAEAVQPRTRFAELGFDLSLTAVELNRL 4817
      ||| | : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      61 SKATGLRLPTLLIFDHTTPAAVARL 86
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : :
Db      4818 NAVTGLRLPPTLVFDHPTPLALSEQL 4843
      : ||||| : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12
 AAY77180
 ID AAY77180 standard; Protein; 12199 AA.
 AC AAY77180,
 DT 05-JUN-2000 (first entry)
 DE S. venezuelae pik gene cluster-encoded protein, SEQ ID NO:6.
 OS Streptomyces venezuelae ATCC15439.
 PN MO200000620-A2.
 PD 06-JAN-2000.
 PF 25-JUN-1999; 99MO-US14398.
 PR 26-JUN-1998; 98US-0105537.
 PA (MINU) UNIV MINNESOTA.
 PY Sherman DH, Liu H, Xue Y, Zhao L,
 DR WPI; 2000-160679/14.
 DR N-PSDB; AA287285.
 XX
 PT Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.
 FT synthesis of methymycin and pikromycin -
 XX
 PS
 XS
 Claim 19; Page 315-353; 438pp; English.
 CC The invention relates to an isolated and purified nucleic acid segment
 CC comprising a desosamine biosynthetic gene cluster, a fragment or its
 CC biologically active variant, where the nucleic acid sequence is not
 CC derived from the eryC gene cluster of *Saccharopolyspora erythraea* or
 CC streptomycis antibioticus. The invention also relates to a macrolide
 CC biosynthetic gene cluster, or fragments thereof. The macrolide
 CC biosynthetic gene cluster encodes proteins which synthesize methymycin,
 CC pikromycin, neomethymycin, nabhomycin or a combination of these
 CC compounds. Recombinant or augmented cells comprising the desosamine
 CC and/or macrolide biosynthetic gene clusters are useful for the production
 CC of biologically active macrolides. The macrolide biosynthetic proteins
 CC are useful for synthesis of methymycin, pikromycin, neomethymycin and
 CC nabhomycin. The alternative termination of polyketide synthesis may be
 CC useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA)
 CC monomers. The compounds produced by the recombinant host cells are useful
 CC as biopolymers, e.g., in packaging or biomedical applications, to
 CC engineer PHA monomer syntheses or to prepare biologically active agents,
 CC such as chemotherapeutics, immunosuppressants, agents to treat asthma,
 CC chronic obstructive pulmonary diseases as well as other diseases involving
 CC respiratory inflammation, cholesterol-lowering agents or macrolide-based
 CC antibiotics which are active against a variety of organisms, e.g.,
 CC bacteria, including multi-drug resistant pneumococci and other

RESULT 14
AA67203
ID AA67203 standard; protein; 1560 AA.
XX
AC AA67203;
XX
DT 23-MAR-2000 (first entry)
XX
DE Narbonolide synthase subunit 3 (PICAIII) protein sequence.
XX
KM Narbonolide polyketide synthase; PKS; narbonolide synthase subunit 3;
KM PICAIII; antibiotic production; narbomycin; picromycin; ketolide.
XX
OS Streptomyces venezuelae.
XX
PN MO9961599-A2.
XX
PD 02-DEC-1999.
XX
PP 27-MAY-1999; 99MO-US11814.
XX
PR 28-MAY-1998; 98US-0087080.
PR 28-AUG-1998; 98US-0141908.
PR 22-SEP-1998; 98US-0100880.
PR 08-FEB-1999; 99US-0119139.
XX
PA (KOSA-) KOSAN BIOSCIENCES INC.
XX
PI Ashley G, Betlach MC, Betlach M, McDaniel R, Tang L;
XX
DR WPI; 2000-072618/06.
DR N-PSDB; AA256001.
XX
PT New recombinant DNA encoding a domain of narbonolide polyketide
synthase, for production of ketolide antibiotics -
XX
PS Example 2; Page 14; 98p; English.
XX
CC This is the Streptomyces venezuelae narbonolide synthase subunit 3,
containing a coding sequence for a narbonolide polyketide synthase
(PKS). Polyketides are compounds synthesized from 2-carbon units through
a series of condensations and subsequent modifications. Modular PKSs are
responsible for the production of many antibiotics including picromycin.
The narbonolide PKS consists of a loading module, six extender modules,
and two thioester domains. Four proteins make up the narbonolide PKS
(PICAI, PICAI, PICAI and PICAI). PICAI includes the loading module
and extender modules 1 and 2, PICAI includes extender modules 3 and 4,
PICAI includes extender module 5 and PICAI includes extender module 6
and a type II thioesterase domain. The second type II thioesterase
domain is found on the PICB protein. The nucleotide sequences encoding
all of these proteins can be isolated in recombinant form from the
recombinant cosmid pKOS023-27 (see AA256001). Narbonolide is
desosaminylated in S. venezuelae to yield narbomycin. The desosamine
transferase enzyme is required for this conversion, and the recombinant
biosynthetic genes are also found in cosmid pKOS023-27. The recombinant
DNA of the invention is used to express, in transformed cells,
narbonolide (or its derivatives) or other ketolides (particularly
hybrids), which may then be converted (e.g. by other enzymes
recombinantly expressed in the same hosts) to polyketide antibiotics or
their intermediates. The antibiotics are useful in human or veterinary
medicine.
XX
SQ Sequence 1560 AA;
XX
Query Match 46.0%; Score 195; DB 21; Length 1560;
Best Local Similarity 48.8%; Pred. No. 2.8e-16;
Matches 42; Conservative 11; Mismatches 29; Indels 4; Gaps 1;
OY 1 LTAISPTQOOSLLDVRATMAVINDGNERATSDAGSPAFHIGFDSVNGVELRNRL 60
DB 1393 ITALPERRPALTLVTRHAAVLAHSHSPDRA---FGRAFTLGSDSLTAVQALNQL 1448

OY 61 SKATGLRPLVTLIFDHTTPAAVARNL 86
DB 1449 STVGNRRLPATTVPDHPFPALAAHL 1474
XX
RESULT 15
AAB18639
ID AAB18639 standard; Protein; 1562 AA.
XX
AC AAB18639;
XX
DT 22-JAN-2001 (first entry)
XX
DE Amino acid sequence of narbonolide synthase subunit 3 (PICAIII).
XX
KM Narbonolide polyketide synthase gene; narbonolide polyketide;
KM antibiotic; C12-hydroxylase; pick; desosamine biosynthesis;
KM desosaminyl transferase enzyme; ketolide; beta-glucosidase enzyme;
KM picromycin biosynthesis.
XX
OS Streptomyces venezuelae.
XX
PN US6117659-A.
XX
PD 12-SEP-2000.
XX
PP 27-MAY-1999; 99US-0320878.
XX
PR 28-MAY-1998; 98US-0087080.
PR 22-SEP-1998; 98US-0100880.
PR 08-FEB-1999; 99US-0119139.
PR 20-MAY-1999; 99US-0134990.
PR 30-APR-1997; 97US-0846247.
PR 06-MAY-1998; 98US-0073538.
PR 28-AUG-1998; 98US-0141908.
XX
PA (KOSA-) KOSAN BIOSCIENCES INC.
XX
PI Ashley G, Betlach MC, Betlach M, Tang L, McDaniel R;
XX
DR WPI; 2000-610844/58.
XX
PT New recombinant pick hydroxylase gene of Streptomyces venezuelae useful
for converting ketolides to antibiotics and as antibiotics and
intermediates in the synthesis of compounds with pharmaceutical value
-
XX
PS Disclosure; Columns 11-14; 117p; English.
XX
CC The present sequence represents a narbonolide synthase subunit 3
(PICAI). The nucleotide sequence encoding it is used in the course of
the invention. The specification describes a recombinant DNA compound
expressing recombinant polyketide synthase genes in host cells for the
production of narbonolide, narbonolide derivatives and polyketides that
are useful as antibiotics and as intermediates in the synthesis of
compounds with pharmaceutical value. The DNA compounds may also encode
a C12-hydroxylase (pick), desosamine biosynthesis and desosaminyl
transferase enzymes (useful for conversion of ketolides to antibiotics),
and the beta-glucosidase enzyme (involved in picromycin biosynthesis).
CC These compounds are also useful for increasing the antibiotic activity
of a compound relative to the unhydroxylated compound. The recombinant
host cells are useful as genetic systems that allow rapid engineering
of the narbonolide polyketide synthase. These would be valuable for
creating novel ketolide analogs for pharmaceutical applications.
XX
SQ Sequence 1562 AA;
XX
Query Match 46.0%; Score 195; DB 21; Length 1562;
Best Local Similarity 48.8%; Pred. No. 2.8e-16;
Matches 42; Conservative 11; Mismatches 29; Indels 4; Gaps 1;
OY 1 LTAISPTQOOSLLDVRATMAVINDGNERATSDAGSPAFHIGFDSVNGVELRNRL 60
DB 1393 ITALPERRPALTLVTRHAAVLAHSHSPDRA---FGRAFTLGSDSLTAVQALNQL 1448

Db	1393	ITAIPEPERRRPALLTVRTHAAAVLGHSSPDRA----	PGRAFTELGFDSLTAVOLRNOL	1448
Qy	61	SKATGLRLPVTLLIPDHTTPAAVAAAL	:	86
Db	1449	STVGNRLPATVFDHPTPAALAAHL	:	1474

Search completed: June 17, 2003, 13:02:28
Job time : 5.21291 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 12:53:02 ; Search time 8.22676 seconds
(without alignments)
4458.180 Million cell updates/sec

Title: US-09-914-286-3_COPY_3537_3714

Perfect score: 882
Sequence: 1 GSVLVYGTGTVGAAVARHL.....LAYRRRAQLPGVSLAMGLM 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	882	100.0	3972	2 Q9S0R8	Q9S0R8 streptomyc
2	882	100.0	5532	2 Q9S0R4	Q9S0R4 streptomyc
3	882	100.0	5532	2 Q9S0R7	Q9S0R7 streptomyc
4	564	63.9	2034	2 Q93NX7	Q93NX7 streptomyc
5	564	63.6	4151	16 Q53490	Q53490 mycobacteri
6	554	62.8	2066	2 Q914X1	Q914X1 streptomyc
7	554	62.8	3654	2 Q30766	Q30766 streptomyc
8	549	62.2	11096	2 Q914M3	Q914M3 streptomyc
9	548.5	62.2	9477	2 Q914X3	Q914X3 streptomyc
10	547	62.0	3739	2 Q92G14	Q92G14 streptomyc
11	546	61.9	4613	2 Q92G15	Q92G15 streptomyc
12	546	61.7	5435	2 Q914X2	Q914X2 streptomyc
13	544	61.0	3613	2 Q93HW6	Q93HW6 streptomyc
14	538	60.9	10917	2 Q93HW6	Q93HW6 streptomyc
15	537	60.9	902	2 Q93AL5	Q93AL5 streptomyc
16	535	60.7	6315	2 Q9ADL6	Q9ADL6 polyangium

17	535	60.7	9507	2 Q9EWA1	Q9EWA1 streptomyc
18	534	60.5	4472	2 Q31954	Q31954 streptomyc
19	532	60.3	5644	2 Q93NX8	Q93NX8 streptomyc
20	531	60.2	4498	2 Q93HE5	Q93HE5 streptomyc
21	526	59.6	1808	2 Q9EWA2	Q9EWA2 streptomyc
22	522	59.2	1807	2 Q93H19	Q93H19 streptomyc
23	521	59.1	3576	16 Q9EX53	Q9EX53 streptomyc
24	520	59.0	3339	2 Q93HJ3	Q93HJ3 streptomyc
25	520	59.0	4735	2 Q54666	Q54666 amycolatops
26	520	59.0	5588	2 Q9ALM2	Q9ALM2 saccharopol
27	519	58.8	8817	2 Q53840	Q53840 polyangium
28	519	58.8	9510	2 Q93NX9	Q93NX9 streptomyc
29	518	58.7	1864	2 Q93955	Q93955 streptomyc
30	512	58.0	1835	2 Q93H85	Q93H85 streptomyc
31	512	58.0	3524	2 Q93H86	Q93H86 streptomyc
32	511	57.9	6797	2 Q9X993	Q9X993 streptomyc
33	509	57.7	2024	2 Q9EWA3	Q9EWA3 streptomyc
34	508	57.6	2152	16 Q9EX54	Q9EX54 streptomyc
35	504.5	57.2	1839	2 Q30765	Q30765 streptomyc
36	504	57.1	4928	2 Q9ALM3	Q9ALM3 saccharopol
37	503.5	57.1	3413	2 Q54593	Q54593 amycolatops
38	500.5	56.7	6048	2 Q93H87	Q93H87 streptomyc
39	496	56.2	1728	2 Q30764	Q30764 streptomyc
40	495.5	56.2	1420	2 Q54591	Q54591 amycolatops
41	493.5	56.0	6146	2 Q93HJ5	Q93HJ5 streptomyc
42	492	55.8	5060	2 Q52545	Q52545 amycolatops
43	492	55.8	5069	2 Q52789	Q52789 amycolatops
44	491	55.7	2731	2 Q923T9	Q923T9 pseudomonas
45	491	55.7	4881	2 Q9S0R3	Q9S0R3 streptomyc

ALIGNMENTS

RESULT 1

Q9S0R8 PRELIMINARY; PRT; 3972 AA.

AC Q9S0R8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 1.
GN AVEA1.
OS Streptomycetes avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9380548; PubMed=10449723;
RA Ikeda H., Nonomiva T., Usami M., Ohta T., Omura S.;
RT "Organization of the biosynthetic gene cluster for the polyketide
antihelminthic macroide avermectin in Streptomycetes avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL; AB032367; BAA84474.1; -
DR InterPro: IPR001227; Ac transferase.
DR InterPro: IPR000794; Ketocacyl-synt.
DR InterPro: IPR003880; Ppantate attach.
DR InterPro: IPR000834; Zn_carpodpept.
DR Pfam; PF00698; Acyl_transf; 3.
DR Pfam; PF02801; ketocacyl-synt_C; 2.
DR Pfam; PF00550; pp-binding; 3_
DR PROSITE; PS00075; ACP_DOMAIN; 3.
DR PROSITE; PS00606; B_KETOCACYL_SYNTHASE; 2.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_2.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 3972 AA; 416852 MW; 2A293695B032B1C3 CRC64;

Query Match 100.0%; Score 882; DB 2; Length 3972;
Best Local Similarity 100.0%; Pred. No. 1.1e-54;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GSVLVTGCTGVLGAAGAAVRLAGVCGVRLDILVSRGPDAPGAEGLRAELALGAEVRIYA 60
DB 3537 GSVLVTGCTGVLGAAGAAVRLAGVCGVRLDILVSRGPDAPGAEGLRAELALGAEVRIYA 3596
QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLLD 120
DB 3597 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLLD 3656
QY 121 ELTRGMEISAFVLFSSAAGILGSAGGNYAANAALDALYRRRAAGLPGVSLANGIM 178
DB 3657 ELTRGMEISAFVLFSSAAGILGSAGGNYAANAALDALYRRRAAGLPGVSLANGIM 3714

RESULT 2
Q9S0R4 PRELIMINARY; PRT; 5532 AA.
AC Q9S0R4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 3.
GN AVEA3.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN (1)
SEQUENCE FROM N.A.
MEDLINE=99380548; PubMed=10449723;
RA Ikeda H., Nonomura T., Usami M., Ohta T., Omura S.;
RT "Organismal macroclide avermectin in Streptomyces avermitilis.";
RT anthelmintic macroclide avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DB EMBL: AB032367; BAA84478.1; -
InterPro: IPR001227; Ac_transferase.
InterPro: IPR000794; ketoacyl-synt.
InterPro: IPR003880; Pantane attach.
InterPro: IPR000834; Zn cardopect.
DR Pfam: PF00698; Acyl_transf. 3.
DR Pfam: PF00109; ketoacyl-synt. 3.
DR Pfam: PF02801; ketoacyl-synt. C; 3.
DR Pfam: PF00550; pp-binding; 3.
DR PROSITE: PS50075; ACP DOMAIN; 3.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 3.
DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 3.
KW Phosphopantetheine; transferase.
SQ SEQUENCE 5532 AA; 575193 MW; EC7B2069D44A057 CRC64;

Query Match 100.0%; Score 882; DB 2; Length 5532;
Best Local Similarity 100.0%; Pred. No.1.6e-54;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSVLVTGCTGVLGAAGAAVRLAGVCGVRLDILVSRGPDAPGAEGLRAELALGAEVRIYA 60
DB 1488 GSVLVTGCTGVLGAAGAAVRLAGVCGVRLDILVSRGPDAPGAEGLRAELALGAEVRIYA 1547
QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLLD 120
DB 1548 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLLD 1607
QY 121 ELTRGMEISAFVLFSSAAGILGSAGGNYAANAALDALYRRRAAGLPGVSLANGIM 178
DB 1608 ELTRGMEISAFVLFSSAAGILGSAGGNYAANAALDALYRRRAAGLPGVSLANGIM 1665

RESULT 3
Q9S0R7 PRELIMINARY; PRT; 6239 AA.
AC Q9S0R7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

```

```

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 2.
GN AVEA2.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN (1)
SEQUENCE FROM N.A.
MEDLINE=99380548; PubMed=10449723;
RA Ikeda H., Nonomura T., Usami M., Ohta T., Omura S.;
RT "Organismal macroclide avermectin in Streptomyces avermitilis.";
RT anthelmintic macroclide avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DB EMBL: AB032367; BAA84475.1; -
InterPro: IPR001227; Ac_transferase.
InterPro: IPR000794; ketoacyl-synt.
InterPro: IPR003880; Pantane attach.
InterPro: IPR000169; SHpol acete.
InterPro: IPR000834; Zn cardopect.
DR Pfam: PF00698; Acyl_transf. 4.
DR Pfam: PF00109; ketoacyl-synt. 4.
DR Pfam: PF02801; ketoacyl-synt. C; 4.
DR Pfam: PF00550; pp-binding; 4.
DR PROSITE: PS50075; ACP DOMAIN; 4.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 4.
DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_4.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 3.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
KW Phosphopantetheine; transferase.
SQ SEQUENCE 6239 AA; 666283 MW; A706280940B502PA CRC64;

Query Match 69.8%; Score 616; DB 2; Length 6239;
Best Local Similarity 69.1%; Pred. No.2.3e-35;
Matches 123; Conservative 16; Mismatches 39; Indels 0; Gaps 0;

QY 1 GSVLVTGCTGVLGAAGAAVRLAGVCGVRLDILVSRGPDAPGAEGLRAELALGAEVRIYA 60
DB 5753 GSVLVTGCTGVLGAAGAAVRLAGVCGVRLDILVSRGPDAPGAEGLRAELALGAEVRIYA 5812
QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLLD 120
DB 5813 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLLD 5872
QY 121 ELTRGMEISAFVLFSSAAGILGSAGGNYAANAALDALYRRRAAGLPGVSLANGIM 178
DB 5873 ELTRGMEISAFVLFSSAAGILGSAGGNYAANAALDALYRRRAAGLPGVSLANGIM 5930

RESULT 4
Q93NX7 PRELIMINARY; PRT; 2034 AA.
AC Q93NX7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE AmpK.
GN AMPK.
OS Streptomyces nodosus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=40318;
RN (1)
SEQUENCE FROM N.A.
RA Caffrey P., Lynch S.V., Flood E.M., Finnan S.M., O'Leary M.;
RT "The amphotericin biosynthetic gene cluster from Streptomyces
RT nodosus.";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DB EMBL: AF357202; AAK73503.1; -
InterPro: IPR001227; Ac_transferase.
InterPro: IPR003880; Pantane short.
InterPro: IPR000794; ketoacyl-synt.
InterPro: IPR003880; Pantane attach.

```

DR InterPro; IPR000379; Ser_estr_sile.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00698; Acyl_transf_1.
DR Pfam; PF00106; adh_short_1.
DR Pfam; PF00109; ketoacyl-synt_1.
DR Pfam; PF02801; ketoacyl-synt_C_1.
DR Pfam; PF00550; pp-binding_1.
DR Pfam; PF00975; Thioesterase; 1.
DR PROSITE; PS00075; ACP_DOMAIN; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
DR Phosphotransferase.
KW SEQUENCE 2034 AA; 212063 MW; 83E962B7DCD5747 CRC64;
Query Match 63.9%; Score 564; DB 2; Length 2034;
Best Local Similarity 65.2%; Pred. No. 3.6e-32;
Matches 116; Conservative 17; Mismatches 45; Indels 0; Gaps 0;
QY 1 GSVLTGTGTGVAARHLAGVCGVRDLTVSRGPDPAAGELRAELALGAERYIVA 60
DB 1386 GTVLTGTGTGGAALRLVTSRGARHLIASRRGPDPAAGTSELVAELTGLGAQVTA 1445
QY 61 CDVGRREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVAKVDAALLD 120
DB 1446 CDVGRDAVDLTVASVPAEPHPLTAVVHTAGVLDALTSGLTPQLAGVLRPADALLH 1505
QY 121 ELTRGMEISAFVLFSSAGILGSAGQGYAANAALDALAYRRRAAGLPGVSLAMGLW 178
DB 1506 EATLQODLAFAFVLYSSISGVIGPGQANYAANASLIDALAHRRRAAGLPGSLAMGPW 1563
RESULT 5
053490 PRELIMINARY; PRT; 4151 AA.
AC 053490;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Polyketide synthase.
GN PK512 OR RV2048C OR MT2108 OR MTV018.35C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigler K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornby T., Jorgensen K., Krogh A., McLaren J., Moule S., Murphy L.,
RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
RA Sultson J.E., Taylor K., Whitehead S., Barrett B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RT Nature 393:537-544 (1998).
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=CDC 1551 / OSHKOSH;
RC Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RT Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL021899; CA117262.1; -;
DR EMBL; AE007061; AKK46387.1; -;
DR HSSP; P25715; IMLA.
DR TIGR; MT2108; -;

DR Tuberculin; RV2048C; -;
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Pantane_attach.
DR Pfam; PF00698; Acyl_transf_2.
DR Pfam; PF00107; adh_zinc_2.
DR Pfam; PF00109; ketoacyl-synt_C_2.
DR Pfam; PF02801; ketoacyl-synt_C_2.
DR Pfam; PF00550; pp-binding_2.
DR PROSITE; PS00075; ACP_DOMAIN; 2.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE; PS00012; PHOSPHOTRANSFERASE; UNKNOWN_2.
KW Phosphotransferase; Transferase; Complete proteome.
FT CONFLICT 2147 2147 H -> Q (IN REF. 2).
FT CONFLICT 2261 2261 DI -> TV (IN REF. 2).
FT CONFLICT 2268 2268 W -> H (IN REF. 2).
FT CONFLICT 2272 2272 S -> A (IN REF. 2).
FT CONFLICT 2279 2280 AF -> PY (IN REF. 2).
FT CONFLICT 2282 2283 AA -> GR (IN REF. 2).
FT CONFLICT 2287 2287 T -> V (IN REF. 2).
FT CONFLICT 2289 2289 F -> W (IN REF. 2).
FT CONFLICT 3004 3004 S -> L (IN REF. 2).
FT CONFLICT 3649 3649 P -> A (IN REF. 2).
SO SEQUENCE 4151 AA; 431577 MW; 53F6F44F11D7A5D CRC64;
Query Match 63.6%; Score 561; DB 16; Length 4151;
Best Local Similarity 60.7%; Pred. No. 1.3e-31;
Matches 108; Conservative 31; Mismatches 39; Indels 0; Gaps 0;
QY 1 GSVLTGTGTGVAARHLAGVCGVRDLTVSRGPDPAAGELRAELALGAERYIVA 60
DB 1680 GTVLTGTGTGGAARHLAGVCGVRDLTVSRGPDPAAGELRAELALGAERYIVA 1739
QY 61 CDVGRREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVAKVDAALLD 120
DB 1740 CDAAARALAKYATIPQHPHLSGVITHAGLDDAVMSLTPTDRVDVLRSKVDAAMH 1799
QY 121 ELTRGMEISAFVLFSSAGILGSAGQGYAANAALDALAYRRRAAGLPGVSLAMGLW 178
DB 1800 ELTRDLDVSAFVLFSSAGILGSAGQGYAANAALDALAYRRRAAGLPGVSLAMGLW 1857
RESULT 6
0914X1 PRELIMINARY; PRT; 2066 AA.
AC 0914X1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE NYSK.
GN Streptomyces noursei.
OS Streptomyces noursei.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=1971;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 11455;
RX MEDLINE=20334850; PubMed=10873841;
RA Brautaset T., Sekurova O.N., Sletta H., Ellingren T.E., Strom A.R.,
RA Valla S., Zotchev S.B.;
RT "Biosynthesis of the polyene antifungal antibiotic nystatin in
RT Streptomyces noursei ATCC 11455: analysis of the gene cluster and
RT deduction of the biosynthetic pathway.";
RT Chem. Biol. 7:395-403 (2000).
DR EMBL; AF263912; AAF71768.1; -;
DR HSSP; P25715; IMLA.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Pantane_attach.

DR InterPro; IPR000379; Ser_estr_ site.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00698; Acyl_transf_1.
 DR Pfam; PF00109; ketoacyl-synt_1.
 DR Pfam; PF02801; ketoacyl-synt_C_1.
 DR Pfam; PF00550; pp-binding_1.
 DR Pfam; PF00975; Thioesterase; 1.
 DR PROSITE; PS50075; ACP DOMAIN; 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 KW Phosphopantetheine; transferase.
 SQ SEQUENCE 206 AA; 214572 MW; C802F5991FB2B56E CRC64;

Query Match 62.8%; Score 554; DB 2; Length 2066;
 Best Local Similarity 64.0%; Pred. No. 1.9e-31;
 Matches 114; Conservative 18; Mismatches 46; Indels 0; Gaps 0;

QY 1 GSVLVTGGTGVGAAGAAVAAHLAGVCGVRDILLVSRGPDAPGAGLPAELALGAEVRIYA 60
 DB 1416 GTVLITGGTGGTGGALALRHLLVTRGARHLLLAGRGPEARGAGELVALTAQADVRVA 1475
 QY 61 CVDGERREVRRLLEGVPACGCLTGVAAGVDDATIASLTPERLGTVPAAKVDALLD 120
 DB 1476 CVDGDTALDALTATVPAAPHLTAVVHTAGVDDALIGSLTPQLATVLRPKADAAWHH 1535
 QY 121 ELTRGMEISAFVLFSSAGILGSAGQGYAANAALDALAYRRRAGLPGVSLAWGLW 178
 DB 1536 DATRGIDLAGFVLYSSVSGVLSFGQGNVAAAANAYDALALARRADGSLPALSIAMQPM 1593

RESULT 7
 ID 030766 PRELIMINARY; PRT; 3654 AA.
 AC 030766;
 DT 01-JUN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Polypeptide synthase modules 4 and 5.
 GN NIDA3.
 OS Streptomyces caelestis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=36816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL-2871;
 RX MEDLINE=98053867; PubMed=9393718;
 RA Kakevab S.O., Katz L., Stasael D.;
 RT "Identification and characterization of the niddamycin polyketide
 RT synthase genes from Streptomyces caelestis.";
 RT J. Bacteriol. 179:7515-7522 (1997).
 DR EMBL; AF016585; AAC46026.1; -
 DR InterPro; IPR001227; Ac_transferase.
 DR InterPro; IPR002085; Adh_zn_family.
 DR InterPro; IPR001827; Antennapedia.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR001993; Mitoch_carrier.
 DR InterPro; IPR003880; Pantne_attach.
 DR InterPro; IPR002364; QOR_zeta_crystal.
 DR Pfam; PF00698; Acyl_transf_2.
 DR Pfam; PF00107; adh_zinc_1.
 DR Pfam; PF00109; ketoacyl-synt_2.
 DR Pfam; PF02801; ketoacyl-synt_C_2.
 DR Pfam; PF00550; pp-binding_2.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 DR PROSITE; PS00032; ANTENNAPEDIA; 1.
 DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 DR PROSITE; PS01162; QOR_ZETA_CRYSTAL; 1.
 KW Phosphopantetheine; transferase.
 SQ SEQUENCE 3654 AA; 379108 MW; 59825E49EDCD06FF CRC64;

Query Match 62.8%; Score 554; DB 2; Length 3654;
 Best Local Similarity 64.6%; Pred. No. 3.6e-31;
 Matches 115; Conservative 17; Mismatches 46; Indels 0; Gaps 0;

QY 1 GSVLVTGGTGVGAAGAAVAAHLAGVCGVRDILLVSRGPDAPGAGLPAELALGAEVRIYA 60
 DB 3192 GTVLITGGTGGTGGALALRHLLVTRGARHLLLAGRGPDGAGLVAELADAGASATVVA 3251
 QY 61 CVDGERREVRRLLEGVPACGCLTGVAAGVDDATIASLTPERLGTVPAAKVDALLD 120
 DB 3252 CVDADADAVRRLIDVEPEQPLTAVVHSGAVLDGMLPSLTPERLERVLRPKVDAVHLD 3311
 QY 121 ELTRGMEISAFVLFSSAGILGSAGQGYAANAALDALAYRRRAGLPGVSLAWGLW 178
 DB 3312 RLTRDLDLSAFVLFSSSALLGSPAGQNYAANAALDALAYRRRAGLPGVSLAWGLW 3369

RESULT 8
 ID 0914W3 PRELIMINARY; PRT; 11096 AA.
 AC 0914W3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Nysc.
 GN NYSO.
 OS Streptomyces noursei.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1971;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 11455;
 RX MEDLINE=20334850; PubMed=10873941;
 RA Brantaset T., Sekurova O.N., Sletta H., Ellingsen T.E., Strom A.R.,
 RA Valla S., Zotchev S.B.;
 RT "Biosynthesis of the polyene antifungal antibiotic nystatin in
 RT Streptomyces noursei ATCC 11455: analysis of the gene cluster and
 RT deduction of the biosynthetic pathway.";
 RT Chem. Biol. 7:395-403 (2000).
 DR EMBL; AF263912; AAF71776.1; -
 DR HSSP; P25715; 1MLA.
 DR InterPro; IPR001227; Ac_transferase.
 DR InterPro; IPR002085; Adh_zn_family.
 DR InterPro; IPR004410; FAD_6.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Pantne_attach.
 DR Pfam; PF00698; Acyl_transf_6.
 DR Pfam; PF00107; adh_zinc_1.
 DR Pfam; PF00109; ketoacyl-synt_6.
 DR Pfam; PF02801; ketoacyl-synt_C_6.
 DR Pfam; PF00550; pp-binding_6.
 DR TIGRfam; TIGR00128; fadD_6.
 DR PROSITE; PS50075; ACP DOMAIN; 6.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 6.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 5.
 KW Phosphopantetheine; transferase.
 SQ SEQUENCE 11096 AA; 1150415 MW; 776CBAFCAE551DD CRC64;

Query Match 62.2%; Score 549; DB 2; Length 11096;
 Best Local Similarity 65.2%; Pred. No. 2.9e-30;
 Matches 116; Conservative 12; Mismatches 50; Indels 0; Gaps 0;

QY 1 GSVLVTGGTGVGAAGAAVAAHLAGVCGVRDILLVSRGPDAPGAGLPAELALGAEVRIYA 60
 DB 8812 GTVLITGGTGGTGGALALRHLLVTRGARHLLLAGRGPDGAGLVAELADAGASATVVA 8871
 QY 61 CVDGERREVRRLLEGVPACGCLTGVAAGVDDATIASLTPERLGTVPAAKVDALLD 120
 DB 8872 CVDAGDAVAAVAVGAPBHPRAVAVHTAGVDDGVISLTIERLATVLRPKADAAWHH 8931
 QY 121 ELTRGMEISAFVLFSSAGILGSAGQGYAANAALDALAYRRRAGLPGVSLAWGLW 178

Db	8932	EATRGDLDDAFVVFSSVAGVFCGAGQANVYAANNAFLDLMQRRRAAGLPGLSLAWGPM	8989
RESULT 9			
Q9L4X3	PRELIMINARY;	PRT, 9477 AA.	
ID	Q9L4X3		
AC	Q9L4X3		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Nysi.		
GN	Nysi.		
OS	Streptomyces noursei.		
OC	Bacteria, Firmicutes, Actinobacteria, Actinobacteridae;		
OC	Actinomycetales, Streptomycineae, Streptomycetaceae, Streptomyces.		
OX	NCBI_TaxID=1971;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 11455;		
FX	MEDLINE=20334650; PubMed=10873841;		
RA	Brattaeac T., Sekurava O.N., Stetta H., Ellingren T.E., Strom A.R.,		
RA	Valla S., Zolcher S.B.;		
RT	"Bioynthesis of the polye antifungal antibiotic mycetanin in		
RT	Streptomyces noursei ATCC 11455: analysis of the gene cluster and		
RT	deduction of the biosynthetic pathway.";		
RL	Chem. Biol. 7:395-403 (2000).		
DR	EMBL, AF263912; AAF71766.1; -.		
DR	HSSP, P81989; 1QES.		
DR	InterPro, IPR001227; Ac transferase.		
DR	InterPro, IPR002198; ADH short.		
DR	InterPro, IPR004410; Fabd.		
DR	InterPro, IPR001899; Gram_pos_anchor.		
DR	InterPro, IPR000794; Ketoacyl-synt.		
DR	InterPro, IPR001005; Myb DNA binding.		
DR	InterPro, IPR003880; Peptide attach.		
DR	Pfam, PF00698; ACYL_transf. 6.		
DR	Pfam, PF00106; adh_short; 3.		
DR	Pfam, PF00109; Ketoacyl-synt. 6.		
DR	Pfam, PF02801; ketoacyl-synt_C; 6.		
DR	Pfam, PF00550; pp-binding; 6.		
DR	TIGRFAMs, TIGR00128; fabd; 6.		
DR	PROSITE, PS00075; ACP DOMAIN; 6.		
DR	PROSITE, PS00606; B_KETOACYL_SYNTHASE; 6.		
DR	PROSITE, PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.		
DR	PROSITE, PS00037; MYB 1; UNKNOWN1.		
DR	PROSITE, PS00012; PHOSPHOPANTETHEINE; 6.		
KW	Phosphopantetheine; Transferase.		
SC	SEQUENCE 9477 AA; 988132 MW; 0D2BCASD7B265483 CRC64;		
Query Match	62.2%; Score 548.5; DB 2; Length 9477;		
Best Local Similarity	63.5%; Pred. No. 2.6e-30;		
Matches 113; Conservative 21; Mismatches 43; Indels 1; Gaps 1;			
Qy	1 GSVLVGTGTGVLGAIVARHLAGVCGVRDLVSRGPRAPAGEGARLALGAERVYA 60		
Db	9050 GVLVTGTGTGALGAIVARRLA-KDAQHILVLSRSGPAPAGAAELRAEIDLGTDTVA 9108		
Qy	61 CDVGRREVRRLLEGVAGCPITGVVHAAGVDDATTIASLTPERLGTVFAAKVDAALLD 120		
Db	9109 CDVADRDQITAVLDLPLADRPITGVVHTAGVDDGVLDRLTERQEVFRAKVTISALLD 9166		
Qy	121 ELTRGMLSAFVLFSSAGILGASAGQGVYAAANALDALAVRRRAAGLPGVSLAWGLW 178		
Db	9169 ELTRDRELAAPVLFSSASAAVGNPGQANVYAANAVLDLALAEQRVLTGIPATSVSGAW 9226		
RESULT 10			
Q9ZG14	PRELIMINARY;	PRT, 3739 AA.	
ID	Q9ZG14		
AC	Q9ZG14		
DT	01-MAY-1999 (TREMBLrel. 10, Created)		
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		

Query	Best Local Similarity	Score	DB 2	Length	3739
Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;					
Qy 1	GSVLVVTGCTGVLGAAVHARHLAGVGVVDLLVSRGPDAGAGRLAEALAGAEVRI	60			
Db 3277	GTVALTGTGCTAGGIVARHVVGVGVRRLVLRGRTDAGAGELVHEALGADSVAA	3336			
Qy 61	CDVGERREVRRLLEGVADGCPRLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLD	120			
Db 3337	CDVADREALTRAVLDSIRAEHRLTRAVVTAGVLDGTLPSMTAEDVBHVLPRKVDAAFLD	3396			
Qy 121	ELTRF-GMELSAFPLPSSAGCIGSAGCGVYAAANALDLARRRAAGIPGVSILAMGLW	178			
Db 3397	ELTSTPEYDLAAFMFSSAAVAFGAGAGAAANATLDLAWRRRTTAGLPALSLGGLW	3456			

```

DR EMBL: AF079136; AAC69329.1; -
DR HSSP: P25715; IMLA.
DR InterPro: IPR001227; AC_transferase.
DR InterPro: IPR000794; ketoacyl-synt.
DR InterPro: IPR003880; Pantane_attach.
DR InterPro: IPR000169; SHprot_acetate.
DR Pfam: PF00698; Acyl_transf_3.
DR Pfam: PF00109; ketoacyl-synt_3.
DR Pfam: PF02801; ketoacyl-synt_C_3.
DR Pfam: PF00550; pp-binding_3.
DR PROSITE: PS50075; ACP_DOMAIN_3.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE_2.
DR PROSITE: PS00012; PHOSPHOPANTHEINE_2.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS; UNKNOMN_1.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 4613 AA; 477281 MW; A3BAF8D37CEC9383 CRC64;

Query Match 61.9%; Score 546; DB 2; Length 4613;
Best Local Similarity 62.8%; Pred. No. 1.8e-30;
Matches 113; Conservative 20; Mismatches 45; Indels 2; Gaps 1;

QY 1 GSVLTGTGTGIVGAAVARRHLAGVCVGRDLLVSRRGPPAPGAEGLRARLALGAEVRIVA 60
DB 4100 GTVLTGTGTGIGLGGVARRHVGEMCVRRLLVSRRGTPAPGADELVHLEALGADVIVA 4159
QY 61 CVYGERREVRRLLEGVPACCPITGVVHAAGVLDATIASLPERLGTVFAAKVDALIID 120
DB 4160 CVADAREALTAVALDAIPEHPLTAAVHAGVLSDTGLPSTMTEDVEHRLRPKVAAFLIID 4219
QY 121 ELTR--GMELSAFVFFSSAAGILGSAGCGNVAANAALDALAYRRRAGLPGVSLAWGL 178
DB 4220 ELTSPADLALFAVNFSSAAVFGAGGAGCAVANAATLIDALAMRRRAAGLPALSLGGLW 4279

RESULT 12
O914X2 PRELIMINARY; PRT: 5435 AA.
AC O914X2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Nysl.
GN NYSL.
OS Streptomyces noursei.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=1971;
RN [1].
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11455;
RC MEDLINE=20334850; PubMed=10873841;
RA Brautaset T., Sekurova O.N., Sletta H., Ellingsen T.E., Strom A.R.,
RA Valla S., Zorchev S.B.;
RT "Biosynthesis of the polyene antifungal antibiotic nystatin in
RT Streptomyces noursei ATCC 11455: analysis of the gene cluster and
RT deduction of the biosynthetic pathway.";
RL Chem. Biol. 7:395-403(2000).
DR EMBL: AF263912; AAF71767.1; -.
DR HSSP: P25715; IMLA.
DR InterPro: IPR001227; AC_transferase.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR002085; Adh_zn_family.
DR InterPro: IPR004410; Pabd.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003880; Pantane_attach.
DR InterPro: IPR003880; Pantane_attach.
DR Pfam: PF00698; Acyl_transf_3.
DR Pfam: PF00106; adh_short_1.
DR Pfam: PF00107; adh_zinc_1.
DR Pfam: PF00109; ketoacyl-synt_3.
DR Pfam: PF02801; ketoacyl-synt_C_3.
DR Pfam: PF00550; pp-binding_3.
DR Pfam: PF00550; pp-binding_3.
DR TIGRfams: TIGR00128; fadb; 3.

```

[illegible]

ID		Q93HU1	PRELIMINARY;	PRR;	3613 AA.
AC	Q93HU1				
DT	01-DEC-2001	(TREMBLrel, 19, Created)			
DT	01-DEC-2001	(TREMBLrel, 19, Last sequence update)			
DT	01-JUN-2002	(TREMBLrel, 21, Last annotation update)			
DE	Modular polyketide synthase.				
GN	OLMA7.				
OS	Streptomyces avermitilis.				
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetaceae;				
OC	Actinomycetales; Streptomyces; Streptomyces;				
OX	NCBI_TaxID=33903;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21477403; PubMed=11572948;				
RA	Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,				
RA	Shinohe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,				
RT	Kikuchi H., Shiba T., Sakaki Y., Hattori M.;				
RT	"Genome sequence of an industrial microorganism Streptomyces				
RT	avermitilis. Deducing the ability of producing secondary				
RT	metabolites."				
RL	Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).				
DR	EMBL; AB070940; BAB69196.1;				
DR	InterPro; IPR001227; Ac transferase.				
DR	InterPro; IPR001298; ADH short.				
DR	InterPro; IPR000794; ketoacyl-synt.				
DR	InterPro; IPR003880; ppanc_n_attach.				
DR	InterPro; IPR001031; Thioesterase.				
DR	Pfam; PF00698; Acyl_transf; 2.				
DR	Pfam; PF00106; adh_short; 1.				
DR	Pfam; PF00109; ketoacyl-synt; 2.				
DR	Pfam; PF02801; ketoacyl-synt_C; 2.				
DR	Pfam; PF00550; pp-binding; 2.				
DR	Pfam; PF00975; Thioesterase; 1.				
DR	PROSITE; PS50075; ACP_DOMAIN; 2.				
DR	PROSITE; PS00606; B KETOACYL SYNTHASE; UNKNOWN_2.				
DR	PROSITE; PS00012; PHOSPHOPANTHEINE; UNKNOWN_2.				
DR	Phosphopantetheine.				
SC	SEQUENCE 3613 AA; 379236 MW; 6FAE5CAB158B442 CRC64;				
QY	Query Match 61.7%; Score 544; DB 2; Length 3613;				
DB	Best Local Similarity 64.0%; Pred. No. 1,9e-30;				
	Matches 114; Conservative 13; Mismatches 51; Indels 0; Gaps 0;				
	1 GSVALTNGTGVGLGAAVARRHLAGCGVARDLLVSRRGDAGAGELRFLAALGAEVIVA 60				
	GVALVVGSTGALGAVARRHVTLAGVRRHVTLAGRGRRAGAAALAEITELGADTRILA 29999				

QY 61 CDVGERREVRLLLEGVPACPLTGVVHAAGVLDATIASLTBRRLGTVPAAKVDALLD 120
 DB 3000 CDVSDRDLALVLLRDIADRPPLTAIVHTAGVLDGTVESLTTPDRPFTVLRAKADAAMHLH 3059
 QY 121 ELTRGHELSAFVLFSSAGILGAGCGNYAANAALDALAYRRRAGLPVGSILAMGLW 178
 DB 3060 ELTODAPLAREFVLFSSAGILGAGCGNYAANAALDALAHRRRAGLPVGSILAMGLW 3117

RESULT 14

ID Q93NM6 PRELIMINARY; PRT; 10917 AA.
 AC Q93NM6; 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE AmphC.
 GN AmphC.
 OS Streptomyces nodosus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxId=40318;
 [1]
 RN SEQUENCE FROM N.A.
 RA Caffrey P., Lynch S.V., Flood E.M., Finan S.M., O'Leary M.;
 RT "The amphotericin biosynthetic gene cluster from Streptomyces
 nodosus";
 RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF357202; AK73514.1; -
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002328; ADH_zinc.
 DR InterPro; IPR002085; Adh zn_family.
 DR InterPro; IPR004410; Fabb.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Ppanene_attach.
 DR Pfam; PF00698; Acyl_transfer; 6.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt; 6.
 DR Pfam; PF02801; ketoacyl-synt; 6.
 DR Pfam; PF00550; pp-binding; 6.
 DR TIGRfam; TIGR00128; Fabb; 6.
 DR PROSITE; PSS0075; ACP_DOMAIN; 6.
 DR PROSITE; PSS0059; ADH_ZINC; UNKNOWN_1.
 DR PROSITE; PSS0060; B_KETOACYL_SYNTHASE; UNKNOWN_6.
 DR PROSITE; PSS0012; PHOSPHOPANTHETHEINE; UNKNOWN_5.
 KW Phosphopantetheine.
 SQ SEQUENCE 10917 AA; 1132905 MW; 15AC5956B85810A1 CRC64;

Query Match 61.0%; Score 538; DB 2; Length 10917;
 Best Local Similarity 63.5%; Pred. No. 1.7e-29;
 Matches 113; Conservative 13; Mismatches 52; Indels 0; Gaps 0;

QY 1 GSVLVGTGTVLGAHAVRHLAGVCGVRDILLVSRGPDAPGAEGLRAELALGAEVRIYA 60
 DB 8695 GTVLVTGGTGGGLGQFARHLVDAVGVNLLVSRGPDAPGAEGLRAELALGAEVRIYA 8754
 QY 61 CDVGERREVRLLLEGVPACPLTGVVHAAGVLDATIASLTBRRLGTVPAAKVDALLD 120
 DB 8755 CDVADPDPAVTAALVAGTAEHRLTAIVHTAGVLDGTVESLTTPDRPFTVLRAKADAAMHLH 8814
 QY 121 ELTRGHELSAFVLFSSAGILGAGCGNYAANAALDALAYRRRAGLPVGSILAMGLW 178
 DB 8815 QATRGDLDAFIYFSSVAGVFGAGGANYAAGNAFLDALMHRRRAGLPVGSILAMGLW 8872

RESULT 15

ID Q93AL5 PRELIMINARY; PRT; 902 AA.
 AC Q93AL5; 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Rearranged polyketide synthase (fragment).

GN NYSCD48.
 OS Streptomyces noursei.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxId=1971;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRD48;
 RA Brautaset T., Bruheim P., Sletta H., Hagen L., Ellingsen T.E.,
 RT "Hexamere derivative of the polyene antifungal antibiotic nystatin is
 produced as a result of an induced rearrangement within the nysC
 polyketide synthase gene in Streptomyces noursei ATCC 11455.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF420594; AL18806.1; -
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Ppanene_attach.
 DR Pfam; PF00109; ketoacyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR PRINTS; PR01610; CD36ANTIGEN.
 DR PROSITE; PSS0075; ACP_DOMAIN; 1.
 DR PROSITE; PSS0012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 FT NON_TER 1
 FT NON_TER 902
 SQ SEQUENCE 902 AA; 93381 MW; C1888F648C12C7CC CRC64;

Query Match 60.9%; Score 537; DB 2; Length 902;
 Best Local Similarity 62.9%; Pred. No. 1.3e-30;
 Matches 112; Conservative 11; Mismatches 55; Indels 0; Gaps 0;

QY 1 GSVLVGTGTVLGAHAVRHLAGVCGVRDILLVSRGPDAPGAEGLRAELALGAEVRIYA 60
 DB 391 GTVLVTGGTGGGLGQFARHLVSRGCGVRHLLTSRSGVGAAGAAGVLELESIGARVVAA 450
 QY 61 CDVGERREVRLLLEGVPACPLTGVVHAAGVLDATIASLTBRRLGTVPAAKVDALLD 120
 DB 451 CDVGDSSAVAEVAVGSESEYPLSAVVAAGVLDGTVESLTTPDRPFTVLRAKADAAMHLH 510
 QY 121 ELTRGHELSAFVLFSSAGILGAGCGNYAANAALDALAYRRRAGLPVGSILAMGLW 178
 DB 511 EATRGDLDAFIYFSSVAGVFGAGGANYAAGNAFLDALMHRRRAGLPVGSILAMGLW 568

Search completed: June 17, 2003, 13:07:58
 Job time : 9.22676 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:52:22 ; Search time 1.9415 Seconds
(without alignments)
3803.313 Million cell updates/sec

Title: US-09-914-286-3_COPY_3537_3714

Perfect score: 882
Sequence: 1 GSVLTGCTGTVLGAIVARHL.....LAVRRRAAGLPGVSLAWGLW 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	475	53.9	3567	1 ERY2_SACER	003132 saccharopol
2	477	48.4	3172	1 ERY3_SACER	003133 saccharopol
3	426.5	48.4	3491	1 ERY1_SACER	003131 saccharopol
4	392	44.4	3519	1 OL56_STPAT	007017 streptomyc
5	295.5	33.5	2110	1 MCAS_MYCO	002251 mycobacteri
6	249	28.2	1774	1 MSAS_PENPA	P22367 penicillium
7	248	28.1	1876	1 PPSB_MYCTU	Q10977 mycobacteri
8	248	28.1	4447	1 PKSK_BACSU	P40803 bacillus su
9	237	26.9	4427	1 PKSK_BACSU	Q05470 bacillus su
10	225.5	25.6	1538	1 PPSB_MYCTU	Q10978 mycobacteri
11	225.5	25.6	2511	1 PAS_CHICK	P12376 gallus gall
12	214.5	24.3	2504	1 PAS_HUMAN	P49327 homo sapien
13	207.5	23.5	2505	1 PAS_RAT	P12785 rattus norv
14	186.5	21.1	838	1 PAS_MOUSE	P19096 mus musculu
15	169	19.2	4273	1 PKSM_BACSU	P40872 bacillus su
16	150.5	17.1	255	1 PABG_MYCAV	P07399 mycobacteri
17	140.5	15.9	255	1 PABG_MYCSM	P71534 mycobacteri
18	135.5	15.4	247	1 PABG_MYCTU	Q48930 mycobacteri
19	134	15.2	251	1 YMP3_STRCO	P43168 streptomyc
20	133.5	15.1	245	1 NODD_RHIS3	P72332 rhizobium s
21	133.5	15.1	247	1 YD50_MYCTU	Q11020 mycobacteri
22	132	15.0	247	1 PABG_PSEBA	Q54438 pseudomonas
23	129	14.6	247	1 YGPF_ECOLI	P52037 escherichia
24	128	14.5	242	1 PABG_ACTAC	P70720 actinobacill
25	125.5	14.2	245	1 NODD_RHIME	P06234 rhizobium m
26	124.5	14.1	320	1 PABG_CUPLA	P28643 cuphea lanc
27	115	13.0	246	1 PHBB_CHRVI	P45375 chromatiu
28	112	12.7	261	1 DHBB_HUMAN	Q92506 homo sapien
29	111.5	12.6	261	1 SPRE_MOUSE	Q64105 mus musculu
30	110.5	12.5	247	1 PABG_SYNY3	P73574 synchocyst
31	110	12.5	241	1 PHBB_RHIME	P50205 rhizobium m
32	110	12.5	241	1 PHBB_ZOORA	P23328 zoogloea ra
33	110	12.5	242	1 PHAB_PARDE	P50204 paracoccus

34	109.5	12.4	246	1 NODD_AZOB	P17611 azospirillum
35	109.5	12.4	260	1 DHBB_MOUSE	P50171 mus musculu
36	109	12.4	736	1 DHBB_HUMAN	P51659 homo sapien
37	105	11.9	248	1 PHAB_ACI5P	P50160 acinetobact
38	104.5	11.8	336	1 T62_MAIZE	P50160 zea mays (m
39	103.5	11.7	262	1 SPRE_RAT	P18297 rattus norv
40	102.5	11.6	244	1 PABG_ECOLI	P25716 escherichia
41	101.5	11.5	247	1 PABG_CHLTR	P38004 chlamydia t
42	101.5	11.5	250	1 LINC_PSEPA	P50197 pseudomonas
43	101	11.5	240	1 PABG_SYNY3	P73826 synchocyst
44	101	11.5	254	1 YJ11_MYCTU	P72057 mycobacteri
45	101	11.5	735	1 DHBB_RAT	P97852 rattus norv

ALIGNMENTS

RESULT 1
ID ERY2_SACER STANDARD; PRT; 3567 AA.
AC 003132; 054096;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Erythronolide synthase, modules 3 and 4 (EC 2.3.1.94) (ORF 2) (6-deoxyerythronolide B synthase II) (DEBS 2).
GN ERYA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Actinobacteria; Actinobacteridae (Class); Actinobacteridae; Actinomycetales; Pseudonocardiales; Pseudonocardaceae;
OC Saccharopolyspora.
OX NCBI_TaxId=1836;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91220065; PubMed=2024119;
RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT "Modular organization of genes required for complex polyketide biosynthesis.";
RT Science 252:675-679 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RX MEDLINE=92155230; PubMed=1740151;
RA Bevil D.J., Cortes J., Haydock S.F., Leadley P.F.;
RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of the structural gene, sequence, analysis and inferred domain structure of the multifunctional enzyme.";
RL Eur. J. Biochem. 204:39-49 (1992).
CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-deoxyerythronolide B.
CC -1- COFACTOR: NADP; CONTAINS 2 COVALENTLY BOUND PHOSPHOTANTHREINES.
CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORF CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERENCE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR) DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.
CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (PAS).
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M63677; AAA26494.1; -.
 CC EMBL; X62569; CAA44449.1; -.
 DR InterPro; IPR001227; Ac. transferase.
 DR InterPro; IPR002085; Adh. zn. family.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantane_attach.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl transfer; 2.
 DR Pfam; PF02801; ketoacyl-synt C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
 DR PROSITE; PS00066; B. KETOACYL SYNTHASE; 2.
 DR PROSITE; PS00075; ACP DOMAIN; 2.
 KW transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KW Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1484
 FT DOMAIN 1485 3567
 FT DOMAIN 27 488
 FT DOMAIN 559 884
 FT DOMAIN 1130 1301
 FT FT
 FT DOMAIN 1397 1467
 FT DOMAIN 1485 1943
 FT DOMAIN 2383 3066
 FT DOMAIN 3139 3322
 FT DOMAIN 3415 3485
 FT ACT SITE 202 202
 FT ACT SITE 651 651
 FT BINDING 1430 1430
 FT ACT SITE 1661 1661
 FT ACT SITE 2115 2115
 FT NP_BIND 2961 2978
 FT NP_BIND 3142 3157
 FT BINDING 3448 3448
 FT CONFLICT 438 438
 FT CONFLICT 480 480
 FT CONFLICT 1241 1241
 FT CONFLICT 2664 2664
 SQ SEQUENCE 3567 AA; 374413 MW; E66284F4738AA0C0 CRC64;

Query Match 53.9%; Score 475; DB 1; Length 3567;
 Best Local Similarity 57.9%; Pred. No. 3.1e-28;
 Matches 103; Conservative 17; Mismatches 56; Indels 2; Gaps 1;

QY 1 GSVLVGTGTVGAARHLAGVGRDILLVSRGPDPAAGARGLAEALGAEVRIVA 60
 DB 3141 GTVLITGGTGTORLRLARHLVTEHGRHLVSRKADAPGSELAEIEDGASAEIAA 3200
 QY 61 CGVGERREVRRLLEGVPACPLTGVVAAAGVLDATIASILTEPERLTGVAKVDALLID 120
 DB 3201 CPTADRDALSLDDGR--PLTGVVAAAGVLDGVTSIDEPAYQVRAKVDAMNH 3258
 QY 121 ELTRGHELSAFVLFSSAAGILSAGCGNTAAANALDALAYRRRAAGLPGVSLAGLW 178
 DB 3259 ELTANTGSLFVLFSSASVLAGPGGVVAAANESINALAALRRTRGLPAKALGMGLW 3316

RESULT 2
 ID ERY3_SACER STANDARD; PRT; 3172 AA.
 AC Q03133; Q54097; Q99270;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Erythronolide synthase, modules 5 and 6 (EC 2.3.1.94) (ORF 3) (6-deoxyerythronolide B synthase III) (DBS 3).
 DE ERYA.

OS Saccharopolyspora erythraea (Streptomyces erythraeus).
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;
 CC Saccharopolyspora.
 CC NCBI_TaxId=1836;
 RX STRAIN=NRRL 2338;
 RX MEDLINE=91043075; PubMed=2234082;
 RX Cortes J., Haydock S.F., Roberts G.A., Bevil D.J., Leadley P.F.;
 RT "An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea.";
 RL Nature 348:176-178 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220065; PubMed=2024119;
 RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
 RT "Molecular organization of genes required for complex polyketide biosynthesis.";
 RL Science 252:675-679 (1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=NRRL 2338;
 RX MEDLINE=92155230; PubMed=1740151;
 RA Bevil D.J., Cortes J., Haydock S.F., Leadley P.F.;
 RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme.";
 RL Eur. J. Biochem. 204:39-49 (1992).
 CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-deoxyerythronolide B.
 CC -1- COFACTOR: NADP; CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.
 CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
 CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
 CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR), DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.
 CC -1- SIMILARITY: TO PARTIAL ACYL SYNTHASE (PAS).
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR	PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
DR	PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR	PROSITE; PS50075; ACQ_DOMAIN; 2.
KW	Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP, Repeat;
KW	Phosphopantetheine; Multifunctional enzyme.
FT	DOMAIN 1 1484
FT	DOMAIN 1485 3172
FT	DOMAIN 37 484
FT	DOMAIN 554 878
FT	DOMAIN 1116 1298
FT	DOMAIN 1394 1464
FT	DOMAIN 1488 1954
FT	DOMAIN 2021 2335
FT	DOMAIN 2555 2735
FT	DOMAIN 2821 2891
FT	DOMAIN 2926 3172
FT	ACT_SITE 199 199
FT	ACT_SITE 643 643
FT	NP_BIND 1118 1164
FT	BINDING 1427 1427
FT	ACT_SITE 1661 1661
FT	ACT_SITE 2112 2112
FT	NP_BIND 2557 2605
FT	BINDING 2854 2854
FT	CONFLICT 231 231
FT	CONFLICT 240 240
FT	CONFLICT 289 289
FT	CONFLICT 493 493
FT	CONFLICT 517 517
FT	CONFLICT 510 510
FT	CONFLICT 513 513
FT	CONFLICT 525 525
FT	CONFLICT 536 536
FT	CONFLICT 547 551
FT	CONFLICT 553 553
FT	CONFLICT 673 673
FT	CONFLICT 716 716
FT	CONFLICT 734 736
FT	CONFLICT 896 896
FT	CONFLICT 896 896
FT	CONFLICT 988 994
FT	CONFLICT 1108 1116
FT	CONFLICT 1124 1126
FT	CONFLICT 1132 1132
FT	CONFLICT 1192 1192
FT	CONFLICT 1194 1194
FT	CONFLICT 1277 1278
FT	CONFLICT 1385 1390
FT	CONFLICT 1485 1485
FT	CONFLICT 1518 1518
FT	CONFLICT 1601 1601
FT	CONFLICT 1724 1725
FT	CONFLICT 1732 1732
FT	CONFLICT 1732 1743
FT	CONFLICT 1762 1762
FT	CONFLICT 2252 2252
FT	CONFLICT 2275 2277
FT	CONFLICT 2408 2408
FT	CONFLICT 2420 2421
FT	CONFLICT 2440 2444
FT	CONFLICT 2596 2596
FT	CONFLICT 2609 2609
FT	CONFLICT 2715 2722
FT	CONFLICT 2754 2754
FT	SEQUENCE 3172 AA; 331474 MW; 331474 DD; 5F CRC64;
DR	PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
DR	PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR	PROSITE; PS50075; ACQ_DOMAIN; 2.
KW	Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP, Repeat;
KW	Phosphopantetheine; Multifunctional enzyme.
FT	DOMAIN 1 1484
FT	DOMAIN 1485 3172
FT	DOMAIN 37 484
FT	DOMAIN 554 878
FT	DOMAIN 1116 1298
FT	DOMAIN 1394 1464
FT	DOMAIN 1488 1954
FT	DOMAIN 2021 2335
FT	DOMAIN 2555 2735
FT	DOMAIN 2821 2891
FT	DOMAIN 2926 3172
FT	ACT_SITE 199 199
FT	ACT_SITE 643 643
FT	NP_BIND 1118 1164
FT	BINDING 1427 1427
FT	ACT_SITE 1661 1661
FT	ACT_SITE 2112 2112
FT	NP_BIND 2557 2605
FT	BINDING 2854 2854
FT	CONFLICT 231 231
FT	CONFLICT 240 240
FT	CONFLICT 289 289
FT	CONFLICT 493 493
FT	CONFLICT 517 517
FT	CONFLICT 510 510
FT	CONFLICT 513 513
FT	CONFLICT 525 525
FT	CONFLICT 536 536
FT	CONFLICT 547 551
FT	CONFLICT 553 553
FT	CONFLICT 673 673
FT	CONFLICT 716 716
FT	CONFLICT 734 736
FT	CONFLICT 896 896
FT	CONFLICT 896 896
FT	CONFLICT 988 994
FT	CONFLICT 1108 1116
FT	CONFLICT 1124 1126
FT	CONFLICT 1132 1132
FT	CONFLICT 1192 1192
FT	CONFLICT 1194 1194
FT	CONFLICT 1277 1278
FT	CONFLICT 1385 1390
FT	CONFLICT 1485 1485
FT	CONFLICT 1518 1518
FT	CONFLICT 1601 1601
FT	CONFLICT 1724 1725
FT	CONFLICT 1732 1732
FT	CONFLICT 1732 1743
FT	CONFLICT 1762 1762
FT	CONFLICT 2252 2252
FT	CONFLICT 2275 2277
FT	CONFLICT 2408 2408
FT	CONFLICT 2420 2421
FT	CONFLICT 2440 2444
FT	CONFLICT 2596 2596
FT	CONFLICT 2609 2609
FT	CONFLICT 2715 2722
FT	CONFLICT 2754 2754
FT	SEQUENCE 3172 AA; 331474 MW; 331474 DD; 5F CRC64;
DR	PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
DR	PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR	PROSITE; PS50075; ACQ_DOMAIN; 2.
KW	Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP, Repeat;
KW	Phosphopantetheine; Multifunctional enzyme.
FT	DOMAIN 1 1484
FT	DOMAIN 1485 3172
FT	DOMAIN 37 484
FT	DOMAIN 554 878
FT	DOMAIN 1116 1298
FT	DOMAIN 1394 1464
FT	DOMAIN 1488 1954
FT	DOMAIN 2021 2335
FT	DOMAIN 2555 2735
FT	DOMAIN 2821 2891
FT	DOMAIN 2926 3172
FT</	

```

Query Match      48.4%; Score 427; DB 1; Length 3172;
Best local similarity 55.9%; Pred. No. 1.2e-24;
Matchee 100; Conservative 17; Mismatches 56; Indels 6; Gaps 4.

0y      1 GSVLTGGTGVLAGAIVARHLAGCCVRDLILYRRGPDPGAEGRRAELTAAGAERVYA 60
|:||||||| | | ||||| | | | | | | | | | | | | | | | | | | | |

```

```

Db 2557 GTALVNGGTALGGHVAHRLAR.-CGVEDLVLSRGVDAPDAAPAAELAEALVALGPKTTITA 26133
Oy 61 CDYGERREVRLLLEGVPA.-GCEPLTGTVHAAGVLDADATIASLTBERLSTVPAAKDALL 119
Db 2616 CDVADREQLSKLLEELRGGORPRTVTHVTAGVSESPLEHIE--GELRSVCAAKRTGARLL 26733
Oy 120 DELTRGHELSAPVLFSSAAGILCSAOCGNVAANAAALDALAARRAAGLGVSLAWGLW 178
Db 2674 DELCPDAE--TFVLFSSGAGVWGSANLGAISAAVAAYDALAHRRRAAGRAATVSAMGAW 2730

RESULT 3
ERYL_SACER
ID ERYL_SACER STANDARD; PRT; 3491 AA.
AC Q03131;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Erythronolide synthase, modules 1 and 2 (EC 2.3.1.94) (ORF 1) (6-deoxyerythronolide B synthase 1) (DBBS 1).
GN ARA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Pseudonocardiales; Pseudonocardiaceae;
OC Saccharopolyspora.
CX NCBI_Taxid=1836;
[1]
RN R.
RP SEQUENCE FROM N.A.
RX MEDLINE=91220065; Pubmed=2024119;
RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT "JSL136, an insertion element in the erythromycin gene, cluster of Saccharopolyspora erythraea.";
RL Gene 126:147-151(1993).
CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-deoxyerythronolide B.
CC -1- COFACTOR: NADP. CONTAINS 3 COVALENTLY BOUND PHOSPHODIANTHETHEINES.
CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
CC -1- BIOSYNTHESIS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNTS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION; BETA-KETOREDUCTASE (KR), DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON- AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.
CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
CC -1- SIMILARITY: CONTAINS 3 ACYL CARRIER DOMAINS.
-----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: M63676; AAA26493.2; -.
DR EMBL: L07626; AAA26504.1; -.
DR HSPD; P25715; IMLA.
DR InterPro: IPR002198; ADH_short.

```

DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantine attach.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl transfer; 3.
 DR Pfam; PF02801; ketoacyl-synt C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
 DR PROSITE; PS50075; ACP_DOMAIN; 3.
 DR Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KW Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1972 MODULE 1.
 FT DOMAIN 1979 3491 ACYLTRANSFERASE (AT) 1.
 FT DOMAIN 1 375 ACYL CARRIER (ACP) 1.
 FT DOMAIN 414 484 BETA-KETOACYL SYNTHASE 1.
 FT DOMAIN 503 961 BETA-KETOACYL REDUCTASE 1.
 FT DOMAIN 1030 1356 ACYLTRANSFERASE (AT) 2.
 FT DOMAIN 1611 1794 BETA-KETOACYL REDUCTASE 1.
 FT DOMAIN 1888 1958 ACYL CARRIER (ACP) 2.
 FT DOMAIN 1979 2441 BETA-KETOACYL SYNTHASE 2.
 FT DOMAIN 2507 2854 ACYLTRANSFERASE (AT) 3.
 FT DOMAIN 3055 3237 BETA-KETOACYL REDUCTASE 2.
 FT DOMAIN 3334 3404 ACYL CARRIER (ACP) 3.
 FT ACT SITE 145 145 ACYL-ENZYME INTERMEDIATE.
 FT ACT SITE 447 447 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT ACT SITE 677 677 BETA-KETOACYL SYNTHASE.
 FT ACT SITE 1128 1128 ACYL-ENZYME INTERMEDIATE.
 FT NP_BIND 1614 1660 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT BINDING 1921 1921 BETA-KETOACYL SYNTHASE.
 FT ACT SITE 2148 2148 ACYL-ENZYME INTERMEDIATE.
 FT ACT SITE 2598 2598 NADP.
 FT NP_BIND 3058 3104 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT BINDING 3367 3367 NADP.
 SQ SEQUENCE 3491 AA; 365022 MW; 682BFC32C90FA8C4 CRC64;

Query Match 48.4%; Score 426.5; DB 1; Length 3491;
 Best Local Similarity 53.9%; Pred. No. 1,4e-24;
 Matches 96; Conservative 18; Mismatches 63; Indels 1; Gaps 1;

QY 1 GSVLTGTCGTCGAVVAHRLAGVCGVRDLVSRGPGAPGAGRLAEALGAEVRIYA 60
 DB 1613 GTVLVTGGTGGVGGQIARLWLR-RGAPHLILVSRGPGDAGELVALEALGARTVA 1671
 QY 61 CDVGERREVRVLLGVPAGCPLTGVVAAAGVLDATLSTPRRLGTVPAKVDALLD 120
 DB 1672 CDVTDBESVARELLGGIGDDVPLSAVFPAATLDDGTVDLTGRIERASRAKVLGARNLH 1731
 QY 121 ELTRGMEISAFVLPSSAAGILGSAGOGNYAANAALDALAYRRRAAGLPVSLAMGLM 178
 DB 1732 ELTRRELDLIRFVLPSSAFASFGAPGLGAYAPGNAVIDGLAQRSDGLPATVAMGTM 1789

RESULT 4
 OL56 STRAT STANDARD; PRT; 3519 AA.
 ID OL56 STRAT 007017;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Oleandomycin polyketide synthase, modules 5 and 6.
 GN ORF.
 OS Streptomyces antibioticus.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1890;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150470; PubMed=8107683;
 RA Swan D.G., Rodriguez A.M., Vilches C., Mendez C., Salas J.A.;
 RT "Characterisation of a Streptomyces antibioticus gene encoding a type I polyketide synthase which has an unusual coding sequence.";

RL Mol. Genet. 242:358-362 (1994).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE BIOSYNTHESIS OF THE OLEANDOMYCIN
 CC LACTONE RING.
 CC -1- CORRELATOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.
 CC -1- SIMILARITY: CONTAINS 2 ACP CARRIER DOMAINS.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: I09654; AAA19695.1; -
 DR HSSP; P25715; 1MLA.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantine attach.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl transfer; 2.
 DR Pfam; PF00975; Thioesterase; 1.
 DR Pfam; PF02801; ketoacyl-synt C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
 DR PROSITE; PS50075; ACP_DOMAIN; 3.
 DR Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;
 KW Phosphopantetheine; Multifunctional enzyme; Repeat.
 FT DOMAIN 1 1972 MODULE 1.
 FT DOMAIN 1979 3491 ACYLTRANSFERASE (AT) 1.
 FT DOMAIN 1 375 ACYL CARRIER (ACP) 1.
 FT DOMAIN 414 484 BETA-KETOACYL SYNTHASE 1.
 FT DOMAIN 503 961 BETA-KETOACYL REDUCTASE 1.
 FT DOMAIN 1030 1356 ACYLTRANSFERASE (AT) 2.
 FT DOMAIN 1611 1794 BETA-KETOACYL REDUCTASE 1.
 FT DOMAIN 1888 1958 ACYL CARRIER (ACP) 2.
 FT DOMAIN 1979 2441 BETA-KETOACYL SYNTHASE 2.
 FT DOMAIN 2507 2854 ACYLTRANSFERASE (AT) 3.
 FT DOMAIN 3055 3237 BETA-KETOACYL REDUCTASE 2.
 FT DOMAIN 3334 3404 ACYL CARRIER (ACP) 3.
 FT ACT SITE 145 145 ACYL-ENZYME INTERMEDIATE.
 FT ACT SITE 447 447 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT ACT SITE 677 677 BETA-KETOACYL SYNTHASE.
 FT ACT SITE 1128 1128 ACYL-ENZYME INTERMEDIATE.
 FT NP_BIND 1614 1660 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT BINDING 1921 1921 BETA-KETOACYL SYNTHASE.
 FT ACT SITE 2148 2148 ACYL-ENZYME INTERMEDIATE.
 FT ACT SITE 2598 2598 NADP.
 FT NP_BIND 3058 3104 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT BINDING 3367 3367 NADP.
 SQ SEQUENCE 3491 AA; 368561 MW; 41AE78AAAE61786 CRC64;

Query Match 44.4%; Score 392; DB 1; Length 3519;
 Best Local Similarity 53.1%; Pred. No. 5.8e-22;
 Matches 95; Conservative 16; Mismatches 64; Indels 4; Gaps 3;

QY 1 GSVLTGTCGTCGAVVAHRLAGVCGVRDLVSRGPGAPGAGRLAEALGAEVRIYA 59
 DB 1202 GTVLVTGGTGGVGGQIARLWLR-RGADHVVLVSRGSGAPGADVRLREGVGARVSR 1260
 QY 60 ACDVGERREVRVLLGVPAGCPLTGVVAAAGVLDATLSTPRRLGTVPAKVDALLD 119
 DB 1261 ACDVADRVARLRLBLLD--GEPTVTVFPAAGVPGSTPLAEISVQAAADVMAKAVGAVYL 1318
 QY 120 DELTRGMEISAFVLPSSAAGILGSAGOGNYAANAALDALAYRRRAAGLPVSLAMGLM 178
 DB 1319 GELVDPGCLGEARFVLPSSNAGVSGGQAVYAANAFLDALAYRRRGGVGLPATVAMGTM 1377

RESULT 5
 MCAS MYCBO STANDARD; PRT; 2110 AA.
 ID MCAS MYCBO


```

AC 002251. 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-1998 (Rel. 36, Last annotation update)
DE Mycroceroic acid synthase.
GN MAS.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinobacteriales;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=92406887; PubMed=1527058;
RA Machur M., Kolatukudy P.E.;
RT "Molecular cloning and sequencing of the gene for mycroceroic acid
RT synthase, a novel fatty acid elongating multifunctional enzyme, from
RT Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guérin,";
RL J. Biol. Chem. 267:19388-19395(1992).
CC -1- FUNCTION: CATALYZES THE ELONGATION OF N-FATTY ACYL-COA WITH
CC METHYLMALONYL-COA (NOT MALONYL-COA) AS THE ELONGATING AGENT TO
CC FORM MYCROCEROYL LIPIDS.
CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETHEINE.
CC -1- SUBUNIT: HOMODIMER WHOSE MONOMERS PROBABLY HAVE A HEAD TO TAIL
CC ARRANGEMENT.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/).
CC or send an email to license@isb-eb.ch).
CC -----
DR EMBL; M95808; AAA5369.1; -.
DR PIR; B44110; B44110.
DR HSSP; P73283; 1ESM.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantane_attach.
DR Pfam; PF00107; adh_zinc.1.
DR Pfam; PF00109; ketoacyl-synt.1.
DR Pfam; PF00550; pp-binding.1.
DR Pfam; PF00698; Acyl_transf.1.
DR Pfam; PF02801; ketoacyl-synt C.1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE.1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE.1.
DR PROSITE; PS50075; ACP_DOMAIN.1.
DR Fatty acid biosynthetase; Multifunctional enzyme; Phosphopantetheine;
KW Transferase; Hydrolase; Oxidoreductase; Ligase; NADP; Membrane.
FT DOMAIN 1 430
FT DOMAIN 533 852
FT DOMAIN ? ?
FT DOMAIN ? ?
FT DOMAIN 2026 2096
FT ACT_SITE 177 177
FT ACT_SITE 623 623
FT NP_BIND 1561 1578
FT NP_BIND 1765 1780
FT BINDING 2059 2059
FT BINDING 2110 AA; 225577 MW; CC658215D7155300 CRC64;
SO SEQUENCE

Query Match 33.5%; Score 295.5; DB 1; Length 2110;
Best Local Similarity 41.3%; Pred. No. 7e-15; Mismatches 66; Indels 13; Gaps 6;
Matches 76; Conservative 29;

1 GSVLTGTTGVLTGAIVARHILGV.-CGVVDLLVSRGPDARCA-----EGLPAELALGAE 55
Db 1764 GSVITITGLGGIGLGFPAKLAAGCG--RIVLTASQGN-PRARQTIGLR----AAQAD 1816
56 VRIYACDVGRREVRVLLIEGCPA.-CGPLTGVVHAAGVLDATATISLTBERIGTVFAKVD 114

```

Db	1817	IVVEBCGNIAPDPADRIVSAANTATGPLERGVLSAAVAEDATLTNTLDELIDRMSKVF	1876
Qy	115	ALLLDELTRGMBLSAFLVFSSAAGTIGSAGCGNYAANAALDALAYRRRAGIGPVSLA	174
Db	1877	GSNMLHRATGQPLDWMCLFSSGALLGSGCGAYAAANWVDVFAHMPAAGLPSVSLA	1936
Qy	175	WGMLM	178
Db	1937	WGAM	1940
RESULT 6			
MSAS_PENPA	STANDARD;	PRT;	1774 AA.
ID	MSAS_PENPA		
AC	P22367;		
DT	01-AUG-1991 (Rel. 19, Created)		
DT	01-AUG-1991 (Rel. 19, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	6-methylsalicylic acid synthase (EC 2.3.1.-) (MSAS).		
OS	Penicillium patulum (Penicillium griseofulvum).		
OC	Eukaryote; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;		
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.		
NCBI_TaxID=5078;			
RP	[1]		
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RC	STRAIN=DSM 62862;		
RA	MEDLINE=91006137; PubMed=2209605;		
RA	Beck J., Ripka S., Siegener A., Schiltz E., Schweizer E.;		
RT	"The multifunctional 6-methylsalicylic acid synthase gene of		
RT	Penicillium patulum. Its gene structure relative to that of other		
RT	polyketide synthases."		
RL	Eur. J. Biochem. 192:487-498(1990).		
CC	-1- FUNCTION: This multifunctional enzyme is a polyketide synthase.		
CC	It catalyzes a total of 11 steps by seven different component		
CC	enzymes, in the biosynthesis of the antibiotic patulin.		
CC	-1- CATALYTIC ACTIVITY: Acetyl-CoA + 3 malonyl-CoA + NADPH =		
CC	-1- 6-methylsalicylic acid + NADP(+) + 3 COA + 3 CO(2) + H(2)O.		
CC	-1- PATHWAY: Patulin biosynthesis.		
CC	-1- SUBUNIT: HOMOMULTIMER.		
CC	-1- INDUCTION: IN THE LATE LOGARITHMIC GROWTH PHASE.		
CC	-1- SIMILARITY: WITH MOST POLYKETIDE SYNTHASES, SPECIALLY WITH RAT		
CC	BATRY ACID SYNTHASE, AND WITH OTHER ENZYMES SUCH AS LIPASES AND		
CC	THIOLASES.		
CC	-----		
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; X55776; CA939295.1; .		
DR	PIR; S13178; S13178.		
DR	InterPro; IPR001227; Ac transferase.		
DR	InterPro; IPR000794; ketoacyl-synt.		
DR	InterPro; IPR003880; Pantane attach.		
DR	Pfam; PF00109; ketoacyl-synt. 1.		
DR	Pfam; PF00550; pp-binding; 1.		
DR	Pfam; PF00658; Acyl_transf; 1.		
DR	Pfam; PF02801; ketoacyl-synt C; 1.		
DR	PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.		
DR	PROSITE; PS50075; B KETOACYL SYNTHASE; 1.		
DR	PROSITE; PS50075; ACP DOMAIN; 1.		
KM	Multifunctional enzyme; Oxidoreductase; Antibiotic biosynthesis;		
KM	transferase; NADP; Phosphopantetheine.		
KM	DOMAIN	186	238
FT	DOMAIN	642	676
FT	DOMAIN	1403	1450
FT	DOMAIN	1700	1769
FT	NP_BIND	1419	1424
FT	ACT_SITE	204	204
FT	ACT_SITE	204	204
FT	BETA-KETOACYL SYNTHASE (BY SIMILARITY).		

FT ACT SITE 653 653 MALONYLTRANSFERASE (BY SIMILARITY).
 FT BINDING 1732 1732 PHOSPHOPANTETHEINE (BY SIMILARITY).
 SQ SEQUENCE 1774 AA; 190732 MW; 0585DD10863P938 CRC64;

Query Match 28.2%; Score 249; DB 1; Length 1774;
 Best Local Similarity 35.7%; Pred. No. 1.9e-11;
 Matches 70; Conservative 26; Mismatches 76; Indels 24; Gaps 6;

QY 1 GSVLTGTVGTVGAAVAARHLAGVGVRLTLVSRGPDAG-----AGRLAEALAL- 52
 DB 1412 GTVLITGGVGLGVLEVADEFLVER-GARRLLISRRA-LPRRTWDVSDLDPTIAKR 1468
 QY 53 -----GAEVRIVACDVERREVRLL-EGVPACPLTGVVHAGVLDATIASLTP 102
 DB 1469 LIESRGASVHVLPLDITKPDVAVQLTALDRLSLPS---VQGVVHAGVLDNELVMQTR 1525
 QY 103 ERIGTVPAKAVDALLDELTRGMELSAFLPSSAGILCSAGCGVYAAANALDALAYR 162
 DB 1526 DAFNRVLAAPKIALALHVEFPPKSYDFVFMSSCGNLVGTQASVSGNAFLDTLATR 1585
 QY 163 RRAAGLPGVSLAWGLW 178
 DB 1586 RARLGDAVSPQWTSW 1601

RESULT 7
 PPSA MYCTU STANDARD; PRT; 1876 AA.

AC Q10977;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 GN Phenolphthiocerol synthesis polyketide synthase ppsa.
 DE PPSA OR RV2931 OR MT3000 OR MTCY338.20.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 [1] _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=H32R;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekle F.,
 RA Badcock K., Baaham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultson J.B., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2] _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey O.,
 RA Kolonay J.F., Nelson W.C., Ormay L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLPHTHIOCEROL
 CC SYNTHESIS.
 CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES (BY
 CC similarity).
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; 274697; CAA98988.1; -
 DR EMBL; AE007122; AAK47328.1; -
 DR TIGR; MT3000; -
 DR TubercuList; RV2931; -
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; ppanine attach.
 DR Pfam; PF00109; ketoacyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl_transf; 1.
 DR Pfam; PF02801; ketoacyl-synt C; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE; PS00075; ACP DOMAIN; 2.
 KW Multifunctional enzyme; Oxidoreductase; Transferase; NADP; Repeat;
 KW Phosphopantetheine; Complete proteome.
 FT DOMAIN 7 80
 FT NP_BIND 1764 1833
 FT NP_BIND 1498 1503
 FT BINDING 62 62
 FT ACT_SITE 273 273
 FT ACT_SITE 720 720
 FT BINDING 1796 1796
 FT CONFLICT 624 624
 FT CONFLICT 877 877
 FT CONFLICT 1323 1323
 SQ SEQUENCE 1876 AA; 198834 MW; D9783DBD48792110 CRC64;

Query Match 28.1%; Score 248; DB 1; Length 1876;
 Best Local Similarity 36.1%; Pred. No. 2.4e-11;
 Matches 69; Conservative 27; Mismatches 77; Indels 18; Gaps 4;

QY 4 LVTVGTGVGAAVAARHLAGVGVRLTLVSRGPDAGAE-----GLRAELAL- 52
 DB 1494 LITGGIGLGLMALMDLDRGHR--LVITGTPLPFRPDMDLDTLRRRIDAIRL 1551
 QY 53 -----GAEVRIVACDVERREVRLL-EGVPACPLTGVVHAGVLDATIASLTPERLCT 107
 DB 1552 ENRGVTVAEVADVCGREDVQALLAARDGAPRIGTHAGTINDQVTSMTGDAVQ 1611
 QY 108 VRAKVDALLDELTRGMELSAFLPSSAGILCSAGCGVYAAANALDALAYRRRAAG 167
 DB 1612 VVMPRTGSGQVLDHAFPPGSDVDFYLTASAGIFGIPGGSYAAANSYDALARARROG 1671
 QY 168 LPGVSLAWGLW 178
 DB 1672 CHTMSLDVYAW 1682

RESULT 8
 PKSX BACSU STANDARD; PRT; 4447 AA.

AC P40803;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 GN Putative polyketide synthase psxk (PSX).
 DE PKSX.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 [1] _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / PB1424;
 RX MEDLINE=95219083; PubMed=7704258;
 RA Albertini A.M., Caramori T., Scoffone F., Scotti C., Gallazzi A.,
 RT "Sequence around the 159 degree region of the Bacillus subtilis
 genome: the psxk locus spans 33.6 kb.";

Microbiology 141:299-309(1995).

SEQUENCE FROM N.A.

STRAIN=168;

MEDLINE=98044033; PubMed=9384377;

Kunert F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertolo M.G., Bessières P., Bolotin A., Borchert S., Bourlès R., Bourlès L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denicot P., Devine K.M., Dueterhoff A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G., Guillep G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Jorls B., Karamata D., Kasehara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Médigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetalle D., Porwollik S., Prescott A.M., Priesgen E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y., Sato T., Scanlan E., Schleich S., Schroeder R., Scoffone F., Sekiguchi J., Sekowski A., Seror S.J., Serron P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Teipstra P., Tognoni A., Tobeo V., Uchiyama S., Vandenbol M., Vannier P., Vassarelli A., Viari A., Wambut R., Wedler E., Wedler H., Welzenegger T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein A.; "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*."

Nature 390:249-256(1997).

-1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN SECONDARY METABOLISM.

-1- COPACTOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES (POTENTIAL).

-1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME FAMILY.

-1- SIMILARITY: CONTAINS 5 ACYL CARRIER DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1sb-rib.ch/announce/> or send an email to license@1sb-rib.ch).

EMBL; U11039; AAA85144.1; -

EMBL; Z59112; CAB13590.1; -

EMBL; Z59113; CAB13601.1; -

HSSP; P14687; 1AMU

Subtilin; BG10930; pK6K.

InterPro; IPR000873; AMP-bind.

InterPro; IPR001242; Condensatn.

InterPro; IPR000794; Ketoacyl-synt.

InterPro; IPR003880; Pantine attach.

Pfam; PF00109; ketoacyl-synt 3.

Pfam; PF00501; AMP-binding; 1.

Pfam; PF00550; pp-binding; 5.

Pfam; PF00668; Condensation; 1.

Pfam; PF02801; ketoacyl-synt C; 3.

PRINTS; PRO0154; AMPBINDING.

PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.

PROSITE; PS00455; AMP BINDING; 1.

PROSITE; PS00606; B-KETOACYL SYNTHASE; 2.

PROSITE; PS00705; ACP DOMAIN; 5.

Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;

Phosphopantetheine; Multifunctional enzyme; Repeat; Ligase;

KW Complete proteome.

FT DOMAIN 1 68 ACYL CARRIER (ACP) 1.

FT DOMAIN 1063 1130 ACYL CARRIER (ACP) 2.

FT DOMAIN 2516 2589 ACYL CARRIER (ACP) 3.

FT DOMAIN 2618 2687 ACYL CARRIER (ACP) 4.

FT DOMAIN 3868 3937 ACYL CARRIER (ACP) 5.

FT BINDING 31 31 PHOSPHOPANTETHEINE (POTENTIAL).

FT BINDING 1093 1093 PHOSPHOPANTETHEINE (POTENTIAL).

FT BINDING 2552 2552 PHOSPHOPANTETHEINE (POTENTIAL).

FT BINDING 2650 2650 PHOSPHOPANTETHEINE (POTENTIAL).

FT ACT SITE 2915 2915 BETA-KETOACYL SYNTHASE (BY SIMILARITY).

FT BINDING 3908 3908 PHOSPHOPANTETHEINE (POTENTIAL).

FT ACT SITE 4147 4147 BETA-KETOACYL SYNTHASE (BY SIMILARITY).

SO SEQUENCE 4447 AA; 496058 MW; 9C6DB4AC0A9C057 CRC64;

Query Match 28.1%; Score 248; DB 1; Length 4447;

Best Local Similarity 36.2%; Pred. No. 5.6e-11;

Matches 67; Conservative 25; Mismatches 85; Indels 8; Gaps 4;

1 GSVLTGNGVGAIVARHLACVCGVRDLVSRGPDPAAGLRRAELALGAEVRIYA 60

2254 GYLLTGGAGSLGLPAKRIARTG-RSTIVTGSVLSSEDENELALRSIGAEVYRE 2312

61 CDVGRREVRVRLLEGVPAGC-PLTGVHAAGVLDATIASLTPERLGTVPAAKVDALL 119

2313 ADVDSQHAVRHLLEIKERYGVLINGIGAGSSCRFFIHKNEPQVLOPKVSGLLHV 2372

120 DELTGMELSAVLFSSAAGILGSGGNYAANAALDALLA-YRRRAGL-----PGVSL 173

2373 DCSQDFPLDFPIFPSSVSGCLGNAGDVAANSPMDAFAYRRSLASKRFGSTISF 2432

174 AWGLW 178

2433 NMPLM 2437

RESULT 9

PKSL_BACSU STANDARD; PRT; 4427 AA.

ID PKSL_BACSU STANDARD; PRT; 4427 AA.

AC 005470;

DT 01-FEB-1995 (rel. 31, Created)

DT 01-FEB-1995 (rel. 31, Last sequence update)

DT 15-JUN-2002 (rel. 41, Last annotation update)

DE Putative polyketide synthase pksl (PKS).

GN PKSL OR PKSX OR PKSA OR OUTG.

OS *Bacillus subtilis*.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OC NCBI_TaxId=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168 / PB1424;

RX MEDLINE=93345624; PubMed=8344529;

RA Scotti C., Piaciti M., Cuzzoni A., Perani P., Tognoni A., Grandi G., Gallizi A., Albertini A.M.,

RT "A *Bacillus subtilis* large ORF coding for a polypeptide highly similar to polyketide synthases."

RT Gene 130:65-71(1993).

RL

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunert F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertolo M.G., Bessières P., Bolotin A., Borchert S., Bourlès R., Bourlès L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denicot P., Devine K.M., Dueterhoff A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G., Guillep G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesli D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudga B., Park S.H.,
 RA Parro V., Pohl T.M., Portelle D., Portolillo S., Prescott A.M.,
 RA Pavesan C., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
 RA Sekiguchi J., Sekowska A., Seror S.O., Serron P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccout E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 RA Toso V., Uchiyama S., Vandenbol M., Vamier F., Vassaretti A.,
 RA Viari A., Wamburt R., Wedler E., Wedler H., Weitzengger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zanchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 [3]
 RP SEQUENCE OF 3619-4427 FROM N.A.
 RC STRAIN=168 / PBI424;
 RA Grandi G.; (JUL-1994) to the EMBL/GenBank/DBJ databases.
 CC Submitted
 CC -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR
 CC THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN
 CC SECONDARY METABOLISM.
 CC -1- COPOLYMER: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES
 CC (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS 5 ACYL CARRIER DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, Z14098; CAI78479.1; --
 CC EMBL, U11033; AAA85145.1; --
 CC EMBL, Z39113; CAB13602.1; --
 CC EMBL, Z35133; CAI84504.1; --
 CC PIR, S25021; S25021.
 CC Subtilist; BG10698; pksL.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Pantane_attach.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00109; ketoacyl-synt; 4.
 DR Pfam; PF00550; pp-binding; 5.
 DR Pfam; PF02801; ketoacyl-synt C; 4.
 DR Prosite; PS00012; PHOSPHOPANTETHEINE; 5.
 DR Prosite; PS00606; B_KETOACYL_SYNTHASE; 1.
 DR Prosite; PS50075; ACP DOMAIN; 5.
 KW Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;
 KW Phosphopantetheine; Multifunctional enzyme; Repeat; Complete proteome.
 FT DOMAIN 211 280
 FT DOMAIN 382 759
 FT DOMAIN 937 1115
 FT DOMAIN 1409 1602
 FT DOMAIN 1687 1759
 FT DOMAIN 1876 2253
 FT DOMAIN 2491 2560
 FT DOMAIN 2632 2701
 FT DOMAIN 2823 3182
 FT DOMAIN 3575 3776
 FT DOMAIN 3854 3923
 FT DOMAIN 4019 4373
 FT BINDING 243 243
 FT BINDING 1723 1723
 FT BINDING 2523 2523

FT BINDING 2664 2664 PHOSPHOPANTETHEINE (POTENTIAL).
 FT BINDING 3886 3886 PHOSPHOPANTETHEINE (POTENTIAL).
 SQ SEQUENCE 4427 AA; 493398 MW; 96125215561AB9F2 CRC64;
 Query Match 26.9%; Score 237; DB 1; Length 4427;
 Best Local Similarity 36.2%; Pred. No. 3.8e-10;
 Matches 67; Conservative 28; Mismatches 82; Indels 8; Gaps 4;
 QY 1 GSVLYVGGVGVGAANARHLAGVCGVRLDLYSRGPPDPAAGAEPAEALGAERVYA 60
 DB 1410 GYLLTGAGAGGJGFPTATIAQNTDAVILGRSPEDRKKKDLA-LQKLGQAIYQ 1468
 QY 61 CVGGERREVRRL-EGVPACCEPLTGVVHAGVLDATIASLTPERLGTVPAAKVDAALL 119
 DB 1469 ADLADQGTVALKKEQNYGSLDGIHISAGLIKONFKKKEVQTVLAPVAGLILH 1528
 QY 120 DEITRGMELSAVLFSSAGILGSAQGVYAAANALDALA-YRRPAGL-----PGVSL 173
 DB 1529 DEATKDIPLDFILFSSGAGAVSAGQADYAMANAFAFSEVRNGQAEHLHRYGKTLISV 1588
 QY 174 ANGLW 178
 DB 1589 CWPLW 1593
 RESULT 10
 PPSE MYCTU STANDARD; PRT; 1538 AA.
 ID PPSE MYCTU
 AC Q10978; 053234;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phenolphthalein synthesis polyketide synthase ppse.
 GN PPSE OR RV2932 OR MT3002 OR MTCY338.21 OR MYV011.01.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_Taxid=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Bignami R., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Bisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLPHICEROL
 CC SYNTHESIS.
 CC -1- COPOLYMER: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTETHEINE (By
 CC similarity).
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use .by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).

CC -----
 CC EMBL AL021070; CA15929.1; -
 CC EMBL AE007122; AAK47329.1; -
 CC TIGR MT3002; -
 CC Tuberculester; RV2932; -
 CC InterPro: IPR001227; Ac transferase.
 CC InterPro: IPR000794; ketoacyl-synt.
 CC InterPro: IPR003880; ppanthe_attach.
 CC Pfam: PF00109; ketoacyl-synt. 1.
 CC Pfam: PF00550; pp-binding; 1.
 CC Pfam: PF00698; Acyl transf; 1.
 CC Pfam: PF02801; ketoacyl-synt C; 1.
 CC PROSITE: PS00606; B_KETOACYL_SINTHASE; 1.
 CC PROSITE: PS00012; PHOSPHOPANTHETHEINE; FALSE_NEG.
 CC PROSITE: PS50075; ACP_DOMAIN; 1.
 CC Hypothetical protein; Multifunctional enzyme; Oxidoreductase;
 CC Transferase; NADP; Phosphopantetheine; Complete proteome.
 CC TRANSFERASE; NADP; Phosphopantetheine (ACQ).
 CC BINDING 1425 1495 ACYL CARRIER (ACP).
 CC SEQUENCE 1538 AA; 162527 MW; B55E2A2042AD00CC CRC64;

Query Match 25.6%; Score 225.5; DB 1; Length 1538;
 Best Local Similarity 35.2%; Pred. No. 1e-09;
 Matches 64; Conservative 32; Mismatches 73; Indels 13; Gaps 5;

QY 2 SVLVGTGTVLGAARHLAGVGVBDLLVSRGPDAPGA-BGLRAELALGAERYIVA 60
 DB 1153 SGLVIGATGIGPHLLRQLARM-GAKTIVAMARK---PALDELQCLAAATGTDLIAVA 1207
 QY 61 CDVGERREVRRLRLE---GVPAACEPLTGVNAAGVLDLDAITATLTBERLCTVAAYVDA 116
 DB 1208 ADATDPAAGTLLDRFETELP---PLEGITLAPAGRPALSSMTDDVTMRPKLDL 1264
 QY 117 LLDLDELTRGMEISAPVLFSSAGILSAGQNYAANAALDALAYRRRAALPGVSLIANG 176
 DB 1265 ALHRSLSKSPVHPVLFSSVSGLSRLMAYTAISAFIDSRAGARTWGLPATVVDNG 1324
 QY 177 LW 178
 DB 1325 LW 1326

RESULT 11
 FAS CHICK STANDARD; PRT; 2511 AA.
 AC P12376;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [includes: EC 2.3.1.38; EC 2.3.1.39;
 EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN OR FAS.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phaethonidae;
 OC Gallus.
 NCBI_TaxID=9031;
 .OX [1]
 RN RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-11.
 RC STRAIN=White Leghorn; TISSUE=Liver;
 RX MEDLINE=95031085; PubMed=7944406;
 RA Huang W.-Y., Chirala S.S., Wakil S.J.;
 RT "Amino-terminal blocking group and sequence of the animal fatty acid
 RT synthase.";
 RL Arch. Biochem. Biophys. 314:45-49(1994).
 RP [2]
 RP SEQUENCE OF 75-1775 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89282777; PubMed=2734291;

RA Holzer K.P., Liu W., Hammes G.G.;
 RT "Molecular cloning and sequencing of chicken liver fatty acid
 RT synthase cDNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4387-4391(1989).
 RN [3]
 RN SEQUENCE OF 1568-2512 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=89139426; PubMed=2917973;
 RA Chirala S.S., Kasturi R., Pazirandeh M., Stolow D.T., Huang W.-Y.,
 RA Wakil S.J.;
 RT "A novel cDNA extension procedure. Isolation of chicken fatty acid
 RT synthase cDNA clones.";
 RL J. Biol. Chem. 264:3750-3757(1989).
 RN [4]
 RN SEQUENCE OF 1752-2512 FROM N.A.
 RX MEDLINE=89320436; PubMed=2842766;
 RA Yuan Z., Liu W., Hammes G.G.;
 RT "Molecular cloning and sequencing of DNA complementary to chicken
 RT liver fatty acid synthase mRNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:6328-6331(1988).
 RN [5]
 RN SEQUENCE OF 2202-2512 FROM N.A.
 RX MEDLINE=89088152; PubMed=3207710;
 RA Kasturi R., Chirala S.S., Pazirandeh M., Wakil S.J.;
 RT "Characterization of a genomic and cDNA clone coding for the
 RT cholesterylase domain and 3' noncoding region of the chicken liver
 RT fatty acid synthase gene.";
 RL Biochemistry 27:7778-7785(1988).
 RN [6]
 RN SEQUENCE OF 2121-2209.
 RX MEDLINE=89192401; PubMed=2648999;
 RA Huang W.-Y., Scoops J.K., Wakil S.J.;
 RT "Complete amino acid sequence of chicken liver acyl carrier protein
 RT derived from the fatty acid synthase.";
 RL Arch. Biochem. Biophys. 270:92-98(1989).
 RN [7]
 RN SEQUENCE OF 2209-2508.
 RC STRAIN=White Leghorn;
 RX MEDLINE=89088151; PubMed=3207709;
 RA Yang C.-Y., Huang W.-Y., Chirala S.S., Wakil S.J.;
 RT "Complete amino acid sequence of the cholesterylase domain of chicken
 RT liver fatty acid synthase.";
 RL Biochemistry 27:7773-7777(1988).
 RN [8]
 RN SEQUENCE OF 667-674 AND 1698-1709.
 RX MEDLINE=89323081; PubMed=2751995;
 RA Chang S.I., Hammes G.G.;
 RT "Amino acid sequences of pyridoxal 5'-phosphate binding sites and
 RT fluorescence resonance energy transfer in chicken liver fatty acid
 RT synthase.";
 RL Biochemistry 28:3781-3788(1989).
 CC [1]- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC [1]- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).
 CC [1]- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC [1]- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC [1]- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC [1]- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NAD(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC [1]- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein]
 CC = 2-hexadecenyl-[acyl-carrier protein] + H(2)O.
 CC [1]- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC [1]- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-
 CC carrier protein] + oleate.
 CC [1]- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC [1]- CAUTION: REF.3 AND REF.5 SEQUENCES DIFFER FROM THAT SHOWN FROM

POSITION 2351 ONWARDS DUE TO A FRAMESHIFT.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: J04485; AAB46389.1; -
 DR EMBL: J03860; AAA46767.1; -
 DR EMBL: J02839; AAA82106.1; ALT_SEQ.
 DR PIR: A33918; XYCHFA.
 DR PIR: A32015; A32015.
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR002085; adh_zn_family.
 DR InterPro: IPR000794; ketoacyl-synt.
 DR InterPro: IPR003880; Ppanne-attach.
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00107; adh_zinc.1.
 DR Pfam: PF00109; ketoacyl-synt.1.
 DR Pfam: PF00550; pp-binding.1.
 DR Pfam: PF00698; Acyl_transfer.1.
 DR Pfam: PF00975; Thioesterase.1.
 DR Pfam: PF02801; ketoacyl-synt C.1.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 1.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE: PS00075; ACP_DOMAIN1.1.
 DR Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 DR Transferrase; Hydrolase; Oxidoreductase; Lyase; NADP; Acetylation;
 DR Alternative splicing; Pyridoxal phosphate.
 FT INIT MET 0
 FT DOMAIN 1 7411
 FT 427 815
 FT 1638 1866
 FT 1867 2119
 FT 2124 2180
 FT DOMAIN 2209 2511
 FT 1 1
 FT MOD RES 160 160
 FT ACT_SITE 579 579
 FT ACT_SITE 877 877
 FT ACT_SITE 1691 1691
 FT 1707 1707
 FT 1888 1903
 FT 2157 2157
 FT 2308 2308
 FT ACT_SITE 2481 2481
 FT 2348 2348
 FT 77 78
 FT 116 116
 FT 675 675
 FT 1169 1169
 FT 1178 1178
 FT 1191 1191
 FT 1198 1198
 FT 1287 1287
 FT 1372 1372
 FT 1533 1533
 FT 1577 1577
 FT 1685 1696
 FT 1732 1732
 FT 1745 1745
 FT 2511 AA; 274648 MM; 622039DAC8315D3F CRC64;
 SQ SEQUENCE

Query Match 25.6%; Score 225.5; DB 1; Length 2511;
 Best Local Similarity 36.2%; Pred. No. 1,6e-09;
 Matches 64; Conservative 26; Mismatches 84; Indels 3; Gaps 2;

Qy 2 SVLVGTGTVGLGAAVNAHLAGVGVDRLLVSRGPDADAGSLRAELAAAGAEVRI VAC 61
 Db 1888 SVITTGGLGFGELGELAAOWMI-ERGAQKVLVTSRSGIRTVQAKCVREWKALGIQVLVRS 1946
 Qy 62 DVGEEREVVLLEGVADGCPPLGVVHAQVLDATIASLTPERLGTVPAPKADALLDE 121
 Db 1947 DVGTELTGTLLEALKLPVGGIFNLAVLADAMIENOTPELFWEVFNPKYSGTLHLD 2006
 Qy 122 LTRGCM-EISAPLTFSSAGILGSGGQVYAAANALDAAVRRRAAGPGVSLANG 176
 Db 2007 VTRKKCPDLDYFVRSVSGRGMAQSNYGFANSMERICGRHHDDGLPLAVQNG 2063
 RESULT 12
 ID_PAS_HUMAN STANDARD; PRT; 2504 AA.
 AC P49327;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [Includes: EC 2.3.1.38; EC 2.3.1.39;
 DE EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN OR FAS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96004605; PubMed=7567999;
 RA Jayakumar A., Tai M.-H., Huang W.-Y., Al-Feel W., Hsu M.,
 RA Abu-Elheiga L., Chitrals S.S., Wakil S.J.;
 RA "human fatty acid synthase: properties and molecular cloning.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:8695-8699 (1995).
 RN [2]
 RP SEQUENCE OF 753-758 AND 1285-1297.
 RX MEDLINE=94294385; PubMed=8022791;
 RA Kuhlida F.P., Jenner K., Wood F.D., Hemmiger R.A., Jacobs L.B.,
 RA Dick J.D., Pasternack G.R.;
 RT "Fatty acid synthase: a potential selective target for
 RT antineoplastic therapy.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6379-6383 (1994).
 CC -FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC -LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC -CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADPH(+).
 CC -CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -CATALYTIC ACTIVITY: (3R)-3-hydroxy-palmitoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC -CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-
 CC carrier protein] + oleate.
 CC -SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC -TISSUE SPECIFICITY: PROMINENT EXPRESSION IN BRAIN, LUNG, AND
 CC LIVER.
 CC -MISCELLANEOUS: THE RELATIVELY LOW BETA-KETOACYL SYNTHASE ACTIVITY
 CC MAY BE ATTRIBUTABLE TO THE LOW 4'-PHOSPHOPANTHETHEINE CONTENT OF
 CC THE PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.embnet.org/announce/> or send an email to license@ebi-embnet.org).

CC EMBL; U26644; AAC50259.1; -
 DR GeneW; HGNC:3594; FASN.
 DR MIM; 600212; -
 DR InterPro; IPR001237; Ac transferase.
 DR InterPro; IPR002085; Adh_zn_family.
 DR InterPro; IPR000794; ketoacyl-lyase.
 DR InterPro; IPR003880; pantoic-acid.
 DR InterPro; IPR000051; SAM bind.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-lyase; 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR Pfam; PF00698; Acyl transferase; 1.
 DR Pfam; PF00975; Thioesterase; 1.
 DR Pfam; PF02801; ketoacyl-lyase; 1.
 DR PROSITE; PS00606; B_KETOACYL-SYNTHASE; 1.
 DR PROSITE; PS00705; ACP DOMAIN; 1.
 KW Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 KW Hydrolyase; Oxidoreductase; Transferase; Lyase; NADP;
 KW Pyridoxal phosphate.
 FT DOMAIN 1 413 BETA-KETOACYL SYNTHASE.
 FT BINDING 428 815 ACYL AND MALONYL TRANSFERASES.
 FT DOMAIN 1630 1857 ENOYL REDUCTASE.
 FT DOMAIN 1858 2113 BETA-KETOACYL REDUCTASE.
 FT DOMAIN 2118 2174 ACYL CARRIER (ACP).
 FT DOMAIN 2202 2504 THIOESTERASE.
 FT ACT_SITE 161 161 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT_SITE 580 580 MALONYLTRANSFERASE (BY SIMILARITY).
 FT ACT_SITE 876 876 BETA-HYDROXYACYL DEHYDRATASE (BY SIMILARITY).
 FT NP_BIND 1666 1683 NADP (ER).
 FT BINDING 1699 1699 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT NP_BIND 1880 1895 NADP (KR).
 FT BINDING 2151 2151 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT ACT_SITE 2302 2302 THIOESTERASE (BY SIMILARITY).
 FT ACT_SITE 2475 2475 THIOESTERASE (BY SIMILARITY).
 SQ SEQUENCE 2504 AA; 273100 MW; 8AAF9809B2338DFA CRC64;
 Query Match 24.3%; Score 214.5; DB 1; Length 2504;
 Best Local Similarity 31.4%; Pred. No. 1.1e-08;
 Matches 64; Conservative 22; Mismatches 61; Indels 57; Gaps 4;
 QY 2 SVLVGTGTGVLGAARHLAGVGVADLLVSRG----- 36
 DB 1880 SYIINGGLGFGLELAQWLI-QRGVQKLVLTSSGIRTVQAKQVRMRGGLQGVST 1938
 QY 37 --PDAFGAGLRAELALGAERIVACDVERREVRLLGCVAGCPPLTGVVAACVLD 94
 DB 1939 NISLEGAAGLLEAALQGL-----PVGVEFLAVLVLD 1971
 QY 95 ATIASLTPLRLGVFAKVDALLDELTRGM--ELSAFVLPSSAGIIGSAGCGVYAA 152
 DB 1972 GLEENOTPEPPOVCKPKTSGLTINLRVREACPELDYFVFFSVSGRGNGMSVYGA 2031
 QY 153 NAALDALAVRRRAAGLPGVSLAWG 176
 DB 2032 NSAMERICKRHRHGGPLGLAVQNG 2055
 RESULT 13
 FAS_RAT STANDARD; PRT; 2505 AA.
 AC P12785; O64717; O09187; O09190;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [includes: EC 2.3.1.38; EC 2.3.1.39;

DE EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Eutezoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxId=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89240686; PubMed=2717611;
 RA Amy C.M., Witkowski A., Naggert J., Williams B., Randhawa Z.,
 RT Smith S.;
 RT "Molecular cloning and sequencing of cDNA encoding the entire rat
 RT fatty acid synthase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3114-3118(1989).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=93075999; PubMed=1319331;
 RA Beck K.F., Schreglmann R., Stachopoulos I., Klein H., Hoch J.,
 RA Schweizer M.;
 RT "The fatty acid synthase (FAS) gene and its promoter in Rattus
 RT norvegicus.";
 RL DNA Seq. 2:359-386(1992).
 RN (3)
 RP SEQUENCE OF 75-2505 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Mammary gland;
 RX MEDLINE=89128431; PubMed=2915923;
 RA Schweizer M., Takabayashi K., Beck K.F., Schreglmann R.,
 RT "Rat mammary gland fatty acid synthase: localization of the
 RT constituent domains and two functional polyadenylation/termination
 RT signals in the cDNA.";
 RL Nucleic Acids Res. 17:567-586(1989).
 RN (4)
 RP SEQUENCE OF 2085-2505 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=88087240; PubMed=2891707;
 RA Naggert J., Witkowski A., Mikkelsen J., Smith S.;
 RT "Molecular cloning and sequencing of a cDNA encoding the thioesterase
 RT domain of the rat fatty acid synthetase.";
 RL J. Biol. Chem. 263:1146-1150(1988).
 RN (5)
 RP SEQUENCE OF 1921-2324 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=87246646; PubMed=3109907;
 RA Witkowski A., Naggert J., Mikkelsen J., Smith S.;
 RT "Molecular cloning and sequencing of a cDNA encoding the acyl carrier
 RT protein and its flanking domains in the mammalian fatty acid
 RT synthetase.";
 RL Eur. J. Biochem. 165:601-606(1987).
 CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxypalmitoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-
 CC carrier protein] + oleate.
 CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M76767; AAA57219.1; -
 DR EMBL; X62888; CAA44679.1; -
 DR EMBL; X62889; CAA44680.1; -
 DR EMBL; X13415; CAA31780.1; -
 DR EMBL; X13527; CAA31882.1; -
 DR EMBL; U03514; AAA41144.1; -
 DR PIR; A30313; XYRTFA.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh zn family.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantane attach.
 DR InterPro; IPR000051; SAM bind.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR Pfam; PF00698; Acyl-transf; 1.
 DR Pfam; PF00975; Thioesterase; 1.
 DR Pfam; PF02801; ketoacyl-synt; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; 1.
 DR PROSITE; PS00075; ACP DOMAIN; 1.
 DR Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 DR Hydroxylase; Oxidoreductase; Transferase; Lyase; NADP;
 DR Pyridoxal phosphate.
 FT DOMAIN 1 413 BETA-KETOACYL SYNTHASE.
 FT DOMAIN 429 817 ACYL AND MALONYL TRANSFERASES.
 FT DOMAIN 1629 1857 ENOYL REDUCTASE.
 FT DOMAIN 1858 2113 BETA-KETOACYL REDUCTASE.
 FT DOMAIN 2118 2174 ACYL CARRIER (ACP).
 FT DOMAIN 2202 2505 THIOESTERASE.
 FT ACT SITE 161 161 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT SITE 581 581 MALONYLTRANSFERASE (BY SIMILARITY).
 FT NP BIND 1665 1682 NADP (ER).
 FT BINDING 1698 1698 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT NP BIND 1765 1780 NADP (KR).
 FT BINDING 1968 1968 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT ACT SITE 2151 2151 THIOESTERASE (BY SIMILARITY).
 FT ACT SITE 2302 2302 THIOESTERASE (BY SIMILARITY).
 FT ACT SITE 2475 2475 BETA-HYDROXYACYL DEHYDRATASE (BY
 FT ACT_SITE 878 878 SIMILARITY).
 FT CONFLICT 871 871 S -> P (IN REF. 3).
 FT CONFLICT 1967 1968 MV -> IL (IN REF. 5).
 FT CONFLICT 2085 2085 C -> P (IN REF. 4).
 FT CONFLICT 2106 2106 A -> V (IN REF. 1 AND 5).
 FT CONFLICT 2296 2296 Y -> H (IN REF. 1 AND 5).
 FT CONFLICT 2505 2505 AA; 272647 MW; 5810EC1D37F3114 CRC64;
 SQ SEQUENCE

Query Match 23.5%; Score 207.5; DB 1; Length 2505;
 Best Local Similarity 34.8%; Pred. No. 3.7e-08;
 Matches 63; Conservative 25; Mismatches 82; Indels 11; Gaps 4;

QY 2 SYLVTCGTVLGAANARHLAAGVGVDDLLVSRGPDAGAGLR-----ELALGAEVR 57
 DB 1880 SYLITGGLGFGIELRLMLV-LRGAORLVLTSSGI-----RTGYQAKHREMRQGIHVL 1934
 QY 58 IVACDGERREVRRLLEGVPAGCELVVHAAGVLDATITASTPERLGTGVPAAKVDAL 117
 DB 1935 VTSNVSLSLEGARALTAATKLGPGVGVRNLAVLDALENQTPPLFDVVKPKNGTL 1994
 QY 118 LDELTRGM--ELSAVLFSSAAGLIGSAGQGVAAANALDALAYRRRAAGLPGVSLAW 175
 DB 1995 NIDRATREACPELDYVAFSSVSCGNGNAGSNYGFANSTMERICGRHHDGLPGLAVQM 2054
 QY 176 G 176

DB 2055 G 2055

RESULT 14
 ID_FAS_MOUSE STANDARD; PRT; 838 AA.
 AC P19096;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [Includes: EC 2.3.1.38; EC 2.3.1.39;
 DE EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14]
 DE (Fragment).
 GN FASN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Liver;
 RX MEDLINE=89149781; PubMed=2920037;
 RA Paulauskis U.D., Sul H.S.;
 RT "Structure of mouse fatty acid synthase mRNA. Identification of the
 RT two NADPH binding sites.";
 RL Biochem. Biophys. Res. Commun. 158:690-695(1989).
 CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC -1- LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC -1- long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
 CC -1- carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxy-palmitoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-
 CC -1- carrier protein] + oleate.
 CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).


```

FT DOMAIN 451 507 ACYL CARRIER (ACP).
FT DOMAIN 535 838 THIOESTERASE.
FT BINDING 484 484 PHOSPHOPANTHETHEINE (BY SIMILARITY).
FT ACT_SITE 635 635 THIOESTERASE (BY SIMILARITY).
FT ACT_SITE 808 808 THIOESTERASE (BY SIMILARITY).
SQ SEQUENCE 838 AA; 91212 MW; 32C40E55CA5AB841 CRC64;

Query Match
Best Local Similarity 33.1%; Score 186.5; DB 1; Length 838;
Matches 60; Conservative 26; Mismatches 84; Indels 11; Gaps 4;

QY 2 SVLVYTGCTGVLGAANVHLAGVCGVRDLLVSRGPDAPGAEGLRA---ELALGAEVR 57
DB 213 SVYITGGLGFGFGLHRLV-LRGAQRLVLTSSRGI-----RFGYQAKHIREWRQGIQVL 267
QY 58 IYACDVGREVRRLLEGVPAAGCPLTGVVHAGVLDATATASLTERLGTVPAAKVDAL 117
DB 268 VSTSNVSLLEGARALLAEATKLGPGVGVNLGHLDAMLENQTPFLFDVKNPKYNGVL 327
QY 118 LLDLRLRGM--ELSAFVLPSSAGILGSAGCGVYAAANALDLAAYRRAPAGLPGVSLAW 175
DB 328 NDRATREACPELDYVAATSSVSCGRGNAGQNTYCFANSTMERICGRHDVLPGLAVOM 387
QY 176 G 176
DB 388 G 388

RESULT 15
PKSM_BACSU STANDARD; PRT; 4273 AA.
ID AC P40872; Q31781;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative polyketide synthase pKsm.
GN PKSM OR PKSY.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN (1)
RN SEQUENCE FROM N.A.
RX MEDLINE=98044033; PubMed=9384377;
RA Kuner F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertsche M.G., Beesler P., Bolochin A., Borchert S.,
RA Borstler R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Broutlet S., Bruch C.V., Caldwell B., Capuano N.J., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denton F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entlin K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,
RA Gilm S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Hartwood C.R., Henaut A.,
RA Hilbert H., Holstappel S., Hosono S., Huille M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaeffer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudena B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Portoullis S., Prescott A.M.,
RA Prescott E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,
RA Sekituchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Taccoti E., Takagi T., Takehashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tostao V., Uchiyama S., Vandenbol M., Vamler F., Vassartotti A.,
RA Viari A., Wambler R., Wedler E., Wedler H., Weizengger T.,
RA Winter P., Wipert A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumelein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus

```

```

RT subtile.";
RL Nature 390:249-256(1997).
RP SEQUENCE OF 1-1763 FROM N.A.
RC STRAIN-168 / PB1424;
RA Tognoni A., Grandi G.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR
CC THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN
CC SECONDARY METABOLISM.
CC -1- CORACTOR: CONTAINS 4 COVALENTLY BOUND PHOSPHOPANTHETHEINES
CC (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 4 ACYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 299113; CAB13603.1; -
DR EMBL; 235133; CA84505.1; -
DR Subtilist; BG10931; pKsm.
DR Interpro: IPR002198; ADH_short.
DR Interpro: IPR000794; Ketoacyl-synt.
DR Interpro: IPR001601; Methyltransf.
DR Interpro: IPR003880; Pentane_attach.
DR Interpro: IPR000051; SAM_bind.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00109; ketoacyl-synt; 3.
DR Pfam; PF00550; pp-binding; 4.
DR PROSITE; PS00801; ketoacyl-synt G; 3.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS50075; ACP_DOMAIN; 4.
KW Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;
KW Phosphopantetheine; Multifunctional enzyme; Repeat; Complete proteome.
FT DOMAIN 295 364
FT DOMAIN 364 4273
FT DOMAIN 364 4273 ACYL CARRIER (ACP) 1.
FT DOMAIN 2190 2258 BETA-KETOACYL SYNTHASE 1.
FT DOMAIN 2190 2258 ACYL CARRIER (ACP) 2.
FT DOMAIN 2322 2737 BETA-KETOACYL SYNTHASE 2.
FT DOMAIN 3532 3947 BETA-KETOACYL SYNTHASE 3.
FT DOMAIN 3410 3483 ACYL CARRIER (ACP) 3.
FT DOMAIN 4140 4209 ACYL CARRIER (ACP) 4.
FT BINDING 327 327 PHOSPHOPANTHETHEINE (POTENTIAL).
FT BINDING 2222 2222 PHOSPHOPANTHETHEINE (POTENTIAL).
FT ACT_SITE 2476 2476 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT BINDING 3446 3446 PHOSPHOPANTHETHEINE (POTENTIAL).
FT ACT_SITE 3690 3690 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT BINDING 4172 4172 PHOSPHOPANTHETHEINE (POTENTIAL).
FT CONFLICT 103 103 E -> V (IN REF. 2).
FT CONFLICT 276 276 Q -> E (IN REF. 2).
FT CONFLICT 289 289 T -> S (IN REF. 2).
SQ SEQUENCE 4273 AA; 477459 MW; 3BBFCFLA250A85A CRC64;

Query Match
Best Local Similarity 29.0%; Score 169; DB 1; Length 4273;
Matches 53; Conservative 27; Mismatches 93; Indels 10; Gaps 4;

QY 1 GSVLYTGCTGVLGAANVHLAGVCGVRDLLVSRGPDAPGAEGLRAELALGAEVRIYA 60
DB 3145 GYVLTGCGVGGVGLFPAHLAGVYAAAN--LITGSPFNDKQKQIKELXDGGERMYAE 3202
QY 61 CVYGERREVRRLLEGVPAAG---CPLTGVVHAGVLDATATASLTERLGTVPAAKVDAA 116
DB 3203 ADVS---DEIANGDCVKGKGRFGAINGVIAHAGIESAIPDKKIESFORIIEPKINGT 3259
QY 117 LLDLRLRGMELSAFVLPSSAGILGSAGCGVYAAANALDLAAYRRAPAGLPGVSLAW 175
DB 3260 IALDELKNEKLDLFWCYSSSSAVLDGDCDYAIGNPOMAYVQIRDELHNGKFTVIMW 3319

```

OY 176 GLW 178
: |
Db 3320 PVW 3322

Search completed: June 17, 2003, 13:02:40
Job time : 2.94115 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 12:55:32 ; Search time 3.97473 Seconds

(without alignments)
4305.183 Million cell updates/sec

Title: US-09-914-286-3_COPY_3537_3714

Sequence: 1 GSVLVTGGTGVGAAGAAVAAHRL.....LAVRRRAAGLPGVSIAMGLW 178

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	561	63.6	4151	2 G70944	probable polyketid
2	547	62.0	3739	2 T17410	polyketide synthase
3	546	61.9	4613	2 T17409	polyketide synthase
4	520	59.0	4735	2 T17463	rifamycin polyketi
5	503.5	57.1	3413	2 T17467	rifamycin polyketi
6	495.5	56.2	1728	2 T17466	rifamycin polyketi
7	482	55.8	5069	2 T17464	rifamycin polyketi
8	488.5	55.4	10223	2 T30225	polyketide synthase
9	482.5	54.7	1616	2 G70668	polyketide synthase
10	481.5	54.6	1937	2 T03224	polyketide synthase
11	481	54.5	6260	2 T30228	polyketide synthase
12	478	54.2	2126	2 H70621	probable polyketid
13	477	54.1	8563	2 T30226	polyketide synthase
14	476.5	54.0	2723	2 T03221	probable polyketid
15	475	53.9	2124	2 T28658	polyketide synthase
16	468.5	53.1	502	2 A70985	probable polyketid
17	467.5	53.0	2103	2 G86925	probable polyketid
18	456.5	51.8	1762	2 T03222	probable polyketid
19	456.5	51.8	2100	2 T03223	probable polyketid
20	450.5	51.1	3573	2 S23070	erythronolide synt
21	433.5	49.1	1763	2 T17465	rifamycin polyketi
22	427	48.4	3491	2 S22012	erythronolide synt
23	426.5	48.2	3491	2 T43231	probable 6-deoxyer
24	425.5	48.2	1562	2 T17411	polyketide synthase
25	403.5	45.7	6420	2 T30283	polyketide synthase
26	402	45.6	3178	2 S13595	6-deoxyerythronol
27	392	44.4	3519	2 S43048	polyketide synthase
28	374.5	42.5	7576	2 T17428	PK506 polyketide s
29	352.5	40.0	2518	2 A12140	polyketide synthase

30	347	39.3	1198	2 T28678	polyketide synthase
31	338	38.3	1570	2 AC2012	hypothetical prote
32	330.5	37.5	2126	2 E70522	probable polyketid
33	328.5	37.2	1582	2 E70876	probable polyketid
34	324.5	36.8	2478	2 A42140	polyketide synthase
35	302.5	34.3	2111	2 A70668	mycocerosate synth
36	299.5	34.0	2108	2 H70819	mycocerosate synth
37	295.5	33.5	2110	2 B44110	mycocerosate synth
38	295.5	33.5	2116	2 C86926	probable mycoceros
39	287.5	32.6	2188	2 A70984	probable polyketid
40	285.5	32.4	2118	2 S72705	mycocerosate synth
41	271.5	30.8	2201	2 S73014	polyketide synthase
42	267	30.3	1827	2 B70984	probable polyketid
43	266	30.2	1774	2 T17421	polyketide synthase
44	265	30.0	958	2 S73012	polyketide synthase
45	265	30.0	1871	2 A87204	polyketide synthase

ALIGNMENTS

RESULT 1
G70944
probable polyketide synthase Rv2048c - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: G70944
R:Colo, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentile, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skellon, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sultson, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; PMID:9829587; PMID:9634230
A:Accession: G70944
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-4151 <COL>
A:Cross-references: GB:AL021899; GB:AL23456; NID:g3242282; PIDN:CA17262.1; PID:g289678
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: pks12
C:Superfamily: Mycobacterium tuberculosis probable polyketide synthase Rv2048c; 3-oxoacyl homology; [acyl-carrier-protein] S-malonyltransferase homology
C:Keywords: carrier protein
F:56-455/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F:559-841/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
F:1680-1861/Domain: short-chain alcohol dehydrogenase homology <SAD1>
F:1963-2038/Domain: acyl carrier protein homology <ACP1>
F:2078-2478/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F:2582-2860/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F:3710-3891/Domain: short-chain alcohol dehydrogenase homology <SAD2>
F:3995-4066/Domain: acyl carrier protein homology <ACP2>

Query Match
Best Local Similarity 63.6%; Score 561; DB 2; Length 4151;
Matches 108; Conservative 31; Mismatches 39; Indels 0; Gaps 0;

QY 1 GSVLVTGGTGVGAAGAAVAAHRLCGVRDLLVSRSGPAPGAGLRPAELALGAERIVA 60
DB 1680 GTVLITGGTGMGSAVARRHVARRHNVIVLSRSGPAPGAAERIVAEILAAAGAQOVVA 1739
QY 61 CDVGRREVRRLLEVPACPLTGVVAAAGVLDATISLTPERIGTYPAARVDAALLD 120
DB 1740 CDADRRAALAKVIADIPVQHPISGVITHAGALDDVVMSTPDRVDVLRSTVDAMHLH 1799
QY 121 ELTRGMEISAFVLFSSAGIGSAGCGVYAAANALDMLAYRRRAAGLPGVSIAMGLW 178
DB 1800 ELTRDLDISAFPMFSSMAGVUSSQANYAAANSTLDLAAHRRAHGLPAISLGMGLW 1857

RESULT 2
T17410

polyketide synthase type I - Streptomyces venezuelae

C:Species: Streptomyces venezuelae

C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000

C:Accession: T17410

R:Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.

Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998

A:Title: A gene cluster for macrocyclic antibiotic biosynthesis in streptomyces venezuelae

A:Reference number: Z18773; PMID:98445333; PMID:9770448

A:Accession: T17410

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3739 <XUE>

A:Cross-references: EMBL:AF079138; NID:G3808326; PID:G3800835; PIDN:AAC69330.1

C:Genetics:

A:Gene: pikA1

C:Superfamily: acyl carrier protein homology

C:Keywords: antibiotic biosynthesis; carrier protein

F:1445-1516/Domain: acyl carrier protein homology <ACP1>

F:3570-3641/Domain: acyl carrier protein homology <ACP2>

Query Match

Best Local Similarity 62.0%; Score 547; DB 2; Length 3739;

Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;

QY 1 GSVLTGCTGTVGAANVAHLAGVCGVRDLTVSRGPDAPGAEGLRAELALGAERVYA 60

DB 3277 GTVLTGCTGCTGALGVARHVGEMVRRLTVSRGPDAPGAEGLVHELEALGADVSA 3336

QY 61 CDVGERREVRLLEGPACPLTGVVHAAGVLDATITASTPERLGTVPFAKVDALLD 120

DB 3337 CDVADREALTAVADISPAHEPLTAVVHTAGVLSDBGLPSMTADVDVHVRPKVDAAFLD 3396

QY 121 ELTR--GMELSAFVLFSAGILGSAGQGYAANAALDALAYRRRAAGLPGVSLAMGLW 178

DB 3397 ELTSTGCTDLAATVMSAAAVFGAGQGYAANAALDALARRKRTAGLPALSLAMGLW 3456

RESULT 3

T17409 polyketide synthase type I - Streptomyces venezuelae

C:Species: Streptomyces venezuelae

C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000

C:Accession: T17409

R:Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.

Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998

A:Title: A gene cluster for macrocyclic antibiotic biosynthesis in streptomyces venezuelae

A:Reference number: Z18773; PMID:98445333; PMID:9770448

A:Accession: T17409

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4613 <XUE>

A:Cross-references: EMBL:AF079138; NID:G3808326; PID:G3800834; PIDN:AAC69329.1

C:Genetics:

A:Gene: pikA1

C:Superfamily: acyl carrier protein homology

C:Keywords: antibiotic biosynthesis; carrier protein

F:1010-1081/Domain: acyl carrier protein homology <ACP1>

F:2495-2566/Domain: acyl carrier protein homology <ACP2>

F:4407-4478/Domain: acyl carrier protein homology <ACP3>

Query Match

Best Local Similarity 61.9%; Score 546; DB 2; Length 4613;

Matches 113; Conservative 20; Mismatches 45; Indels 2; Gaps 1;

QY 1 GSVLTGCTGTVGAANVAHLAGVCGVRDLTVSRGPDAPGAEGLRAELALGAERVYA 60

DB 4100 GTVLTGCTGCTGALGVARHVGEMVRRLTVSRGPDAPGAEGLVHELEALGADVSA 4159

QY 61 CDVGERREVRLLEGPACPLTGVVHAAGVLDATITASTPERLGTVPFAKVDALLD 120

DB 4160 CDVADREALTAVADISPAHEPLTAVVHTAGVLSDBGLPSMTADVDVHVRPKVDAAFLD 4219

QY 121 ELTR--GMELSAFVLFSAGILGSAGQGYAANAALDALAYRRRAAGLPGVSLAMGLW 178

Db 4220 ELTSTGCTDLAATVMSAAAVFGAGQGYAANAALDALARRKRTAGLPALSLAMGLW 4279

RESULT 4

T17463 rifamycin polyketide synthase modules 1-3 - Amycolatopsis mediterranei

C:Species: Amycolatopsis mediterranei

C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000

C:Accession: T17463

R:Schupp, T.

submitted to the EMBL Data Library, December 1997

A:Reference number: Z18802

A:Accession: T17463

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4735 <SCH>

A:Cross-references: EMBL:AJ223012; NID:e1227119; PID:e1227120; PIDN:CAA11035.1

A:Experimental source: strain LBG A3136

C:Superfamily: acyl carrier protein homology; acetate-CoA ligase homology

C:Keywords: carrier protein

F:53-500/Domain: acetate-CoA ligase homology <ACL>

F:543-610/Domain: acyl carrier protein homology <ACP1>

F:2102-2173/Domain: acyl carrier protein homology <ACP2>

F:3079-3150/Domain: acyl carrier protein homology <ACP3>

F:4578-4649/Domain: acyl carrier protein homology <ACP4>

Query Match

Best Local Similarity 59.0%; Score 520; DB 2; Length 4735;

Matches 109; Conservative 20; Mismatches 45; Indels 4; Gaps 2;

QY 1 GSVLTGCTGTVGAANVAHLAGVCGVRDLTVSRGPDAPGAEGLRAELALGAERVYA 60

DB 1829 GTVLTGCTGCTGALGVARHVGEMVRRLTVSRGPDAPGAEGLVHELEALGADVSA 1884

QY 61 CDVGERREVRLLEGPACPLTGVVHAAGVLDATITASTPERLGTVPFAKVDALLD 120

DB 1885 CDVADREALTAVADISPAHEPLTAVVHTAGVLSDBGLPSMTADVDVHVRPKVDAAFLD 1944

QY 121 ELTR--GMELSAFVLFSAGILGSAGQGYAANAALDALAYRRRAAGLPGVSLAMGLW 178

DB 1945 ELTSTGCTDLAATVMSAAAVFGAGQGYAANAALDALARRKRTAGLPALSLAMGLW 2002

RESULT 5

T17467 rifamycin polyketide synthase modules 9-10 - Amycolatopsis mediterranei

C:Species: Amycolatopsis mediterranei

C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000

C:Accession: T17467

R:Schupp, T.

submitted to the EMBL Data Library, December 1997

A:Reference number: Z18802

A:Accession: T17467

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3413 <SCH>

A:Cross-references: EMBL:AJ223012; NID:e1227119; PID:e1227124; PIDN:CAA11039.1

A:Experimental source: strain LBG A3136

C:Superfamily: acyl carrier protein homology

C:Keywords: carrier protein

F:1608-1679/Domain: acyl carrier protein homology <ACP1>

F:3334-3405/Domain: acyl carrier protein homology <ACP2>

Query Match

Best Local Similarity 57.1%; Score 503.5; DB 2; Length 3413;

Matches 113; Conservative 14; Mismatches 45; Indels 7; Gaps 3;

QY 1 GSVLTGCTGTVGAANVAHLAGVCGVRDLTVSRGPDAPGAEGLRAELALGAERVYA 60

DB 1338 GTVLTGCTGCTGALGVARHVGEMVRRLTVSRGPDAPGAEGLVHELEALGADVSA 1396

QY 61 CDVGERREVRLLEGPACPLTGVVHAAGVLDATITASTPERLGTVPFAKVDALLD 119

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: G70668
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1616 <COL>
 A:Cross-references: GB:283658; GB:AL123456; NID:g3261675; PID:CA806103.1; PID:g1781167
 A:Experimental source: strain H37RV
 A:Smith, D.R.; Robison, K.
 submitted to the EMBL Data Library, September 1994
 A:Description: *Mycobacterium tuberculosis* cosmid tbc2.
 A:Reference number: S73053
 A:Accession: S73075
 A:Molecule type: DNA
 A:Residues: 'MMGARMVACLSSTEL', 48-917, 'RSEARSRRGCMGSRAPAPPGIPGIP', 'PMPTPIYERAPGVFN
 A:Cross-references: EMBL:U00024; NID:g560506; PID:AAA50928.1; PID:g560507
 A:Genetics:
 A:Gene: pks1
 A:Superfamily: *Mycobacterium tuberculosis* probable polyketide synthase pks1; acyl carrier
 ology
 C:Keywords: carrier protein
 F:84-162/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
 F:1228-1410/Domain: short-chain alcohol dehydrogenase homology <SADH>
 F:1514-1585/Domain: acyl carrier protein homology <ACPI>

Query Match 54.7%; Score 482.5; DB 2; Length 1616;
 Best Local Similarity 56.4%; Pred. No. 5.3e-30;
 Matches 101; Conservative 23; Mismatches 54; Indels 1; Gaps 1;

QY 1 GSVLVGTGTVGAANVAHLAGVCGVRDLVSRGPDAPGAEGLRAELALGAERYIA 60
 DB 1228 GTVITGATGAGVGLARHLVGAIVGRHLVLASRRDRAEGAEALADTEGAKYQVVA 1287
 QY 61 CDVGERREVRLLLEGVPAC-PLTVVNAAGVLDATIASLTPERLGTVPAAKVDAALL 119
 DB 1288 CVVADAAVAGLPAQISREYPPRGVIAHAGVLDVITSLTPDRIDTVLRKVDAAMWL 1347
 QY 120 DELTRGMEISAPVLFSSAGILGASAGQVYAAANALDLAARRRANGPGVSLANGWL 178
 DB 1348 HQATSLDLSMPLALCSIAATYSGPQGNYSANANFLDGLAHRQAGLAGISLAWGLW 1406

RESULT 10

T30224

probable polyketide synthase module 4 - Streptomyces hygroscopicus
 C:Species: Streptomyces hygroscopicus
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 26-May-2000
 C:Accession: T30224
 R:Ruan, X.; Staessli, D.; Lax, S.; Katz, L.
 Gene 203, 1-9, 1997

A:Title: A second type-I PKS gene cluster isolated from Streptomyces hygroscopicus ATCC
 A:Reference number: Z14848; MUID:98085969; PMID:9426000
 A:Accession: T30224
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1937 <RNA>
 A:Cross-references: EMBL:AF007101; NID:g2624946; PID:AA308064.1; PID:g2624951

A:Experimental source: ATCC 29253
 C:Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoac
 homology; [acyl-carrier-protein] S-malonyltransferase homology
 C:Keywords: carrier protein
 F:55-456/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
 F:549-827/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>

Query Match 54.6%; Score 481.5; DB 2; Length 1937;
 Best Local Similarity 57.3%; Pred. No. 7.7e-30;
 Matches 102; Conservative 19; Mismatches 54; Indels 3; Gaps 1;

QY 1 GSVLVGTGTVGAANVAHLAGVCGVRDLVSRGPDAPGAEGLRAELALGAERYIA 60
 DB 1319 GTTLITGGTGSGLVARRHLVTEHGVRLVLSRQPDADSTELAAELQHGARRVTIS 1378

QY 61 CDVGERREVRLLLEGVPACPLTGVVNAAGVLDATIASLTPERLGTVPAAKVDAALLD 120
 DB 1379 COLTRTALATLIDEG---PLTVVNAAGALADATVHELDALATTAPAAKVAAMWL 1435

QY 121 ELTRGMEISAPVLFSSAGILGASAGQVYAAANALDLAARRRANGPGVSLANGWL 178
 DB 1436 ELTQDHDLTLFALFSSVAGILGSPQGNYSANANFLDGLAHRQAGLAGISLAWGWS 1493

RESULT 11

T30228

Polyketide synthase - Streptomyces hygroscopicus
 C:Species: Streptomyces hygroscopicus
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
 C:Accession: T30228

R:Apaticio, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Stoun
 Gene 169, 9-16, 1996
 A:Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hy
 A:Reference number: Z20782; MUID:96186896; PMID:8635756
 A:Accession: T30228

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-6260 <APA>
 A:Cross-references: EMBL:X86780; NID:g987088; PID:g987102; PID:CAA60462.1
 A:Genetics:
 A:Gene: rapC
 C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protei

C:Keywords: carrier protein
 F:55-451/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
 F:1536-1507/Domain: acyl carrier protein homology <ACPI>
 F:1536-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
 F:3123-3194/Domain: acyl carrier protein homology <ACP2>
 F:3123-3632/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
 F:3725-3999/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
 F:5114-5185/Domain: acyl carrier protein homology <ACP3>
 F:5242-5638/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
 F:5759-6044/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
 F:1135-6206/Domain: acyl carrier protein homology <ACP4>

Query Match 54.5%; Score 481; DB 2; Length 6260;
 Best Local Similarity 56.7%; Pred. No. 2.8e-29;
 Matches 101; Conservative 21; Mismatches 52; Indels 4; Gaps 1;

QY 1 GSVLVGTGTVGAANVAHLAGVCGVRDLVSRGPDAPGAEGLRAELALGAERYIA 60
 DB 1270 GTVLITGSSGVLAGAARHLVTERGVRLHLLSRAP---EALIGELGELGAEVETRA 1325

QY 61 CDVGERREVRLLLEGVPACPLTGVVNAAGVLDATIASLTPERLGTVPAAKVDAALLD 120
 DB 1326 CDVSDPALITQVLAVSPSEHPLTAVIHTAGVDDGVESLTFQRLLETVLRPADGAMWL 1385

QY 121 ELTRGMEISAPVLFSSAGILGASAGQVYAAANALDLAARRRANGPGVSLANGWL 178
 DB 1386 ELTRDADLAAFMVSSAGVLSAGQVYAAANALDLAARRRANGPGVSLANGWL 1443

RESULT 12

H70621

probable polyketide synthase - *Mycobacterium tuberculosis* (strain H37RV)
 C:Species: *Mycobacterium tuberculosis*
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: H70621

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
 ; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Sgares, S.
 Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: H70621

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-2126 <COL>

F:370/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 53.9%; Score 475; DB 2; Length 2124;

Best Local Similarity 55.1%; Pred. No. 2.7e-29;

Matches 98; Conservative 27; Mismatches 53; Indels 0; Gaps 0;

```
QY 1 GSVLVTTGGTGVGAANAARHLAGVCGVRDLLVSRGPDAPGAEGLRAELALGAEVRIYA 60
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1819 GTVLITGGTGTGLGALVARRLVNNHDAKHLLTSRQGSAPGADVLRSELEALGASVTIAA 1878
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 CDYGERREVVRLLGVPAGCPLTGVVHAAGVLDATITASTPERLGTVPFAKVDALLLD 120
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1879 CDVADPRALKDLDNIIPSAHPVAAYVPAASVLDGDLGAMSLERIDRVFAPKIDAAWHLH 1938
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 121 ELTRGMELSAFVLFFSSAGILGSAGCGNVAANAALDALYRRRAAGLPGVSLAMGLW 178
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1939 QLTQDKPLAAFIIFSSVAGVLSGSGHSNYPAAASAFLDALAHRRRAQGLPASSLAMSHW 1996
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

Search completed: June 17, 2003, 13:12:25
Job time : 4.97473 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using ew model

Run on: June 17, 2003, 13:08:02 ; Search time 5.54613 Seconds
(without alignments)
3431.399 Million cell updates/sec

Title: US-09-914-286-3_COPY_3537_3714
Perfect score: 882
Sequence: 1 GSVLVTGCTGVGAARH.....LAYRRRAAGLPGVSLAMGLW 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqe, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Published Applications_A1*
2: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	547	62.0	3739	9	US-09-860-846-33 Sequence 33, Appl
2	547	62.0	3739	9	US-09-988-384B-33 Sequence 33, Appl
3	547	62.0	3739	9	US-09-836-821-33 Sequence 33, Appl
4	547	62.0	3739	9	US-09-793-708-2 Sequence 2, Appl
5	547	62.0	3739	10	US-09-861-289-33 Sequence 33, Appl
6	547	62.0	11877	9	US-09-860-846-6 Sequence 6, Appl
7	547	62.0	11877	9	US-09-836-821-6 Sequence 6, Appl
8	547	62.0	11877	10	US-09-861-289-6 Sequence 6, Appl
9	547	62.0	12199	9	US-09-988-384B-6 Sequence 6, Appl
10	546	61.9	4551	9	US-09-793-708-1 Sequence 1, Appl
11	546	61.9	4613	9	US-09-860-846-31 Sequence 31, Appl
12	546	61.9	4613	9	US-09-988-384B-31 Sequence 31, Appl
13	546	61.9	4613	9	US-09-836-821-31 Sequence 31, Appl
14	546	61.9	4613	10	US-09-861-289-31 Sequence 31, Appl
15	530	60.1	5215	9	US-09-860-846-2 Sequence 2, Appl
16	530	60.1	5215	9	US-09-988-384B-2 Sequence 2, Appl
17	530	60.1	5215	9	US-09-836-821-2 Sequence 2, Appl
18	530	60.1	5215	10	US-09-861-289-2 Sequence 2, Appl
19	482.5	54.7	1616	9	US-09-712-363-262 Sequence 262, App

20	449.5	51.0	3816	9	US-09-808-880-3 Sequence 3, Appl
21	440.5	49.9	7257	9	US-10-014-717-5 Sequence 5, Appl
22	428	48.5	4150	9	US-09-808-880-2 Sequence 2, Appl
23	425.5	48.2	1562	9	US-09-860-846-35 Sequence 35, Appl
24	425.5	48.2	1562	9	US-09-988-384B-35 Sequence 35, Appl
25	425.5	48.2	1562	9	US-09-836-821-35 Sequence 35, Appl
26	425.5	48.2	1562	9	US-09-793-708-3 Sequence 3, Appl
27	425.5	48.2	1562	10	US-09-861-289-35 Sequence 35, Appl
28	395	44.8	184	9	US-09-957-483-33 Sequence 33, Appl
29	392	44.4	3519	9	US-09-808-880-4 Sequence 4, Appl
30	391	44.3	2439	9	US-10-014-717-7 Sequence 7, Appl
31	381	43.2	184	9	US-09-957-483-35 Sequence 35, Appl
32	380	43.1	1832	9	US-10-014-717-4 Sequence 4, Appl
33	336.5	38.2	3798	9	US-10-014-717-6 Sequence 6, Appl
34	267	30.3	1827	9	US-09-712-363-261 Sequence 261, App
35	241.5	27.4	2563	9	US-09-836-705-46 Sequence 46, Appl
36	214.5	24.3	3209	9	US-10-237-271-1 Sequence 1, Appl
37	196	22.2	3032	9	US-09-836-705-44 Sequence 44, Appl
38	177	20.1	711	9	US-10-124-800-18 Sequence 18, Appl
39	177	20.1	1481	9	US-10-331-061-70 Sequence 70, Appl
40	177	20.1	2910	9	US-10-124-800-2 Sequence 2, Appl
41	165.5	18.8	2756	9	US-10-331-061-7 Sequence 7, Appl
42	132	15.0	247	10	US-09-815-242-11893 Sequence 11893, A
43	122.5	13.9	243	10	US-09-815-242-13360 Sequence 13360, A
44	122.5	13.9	243	10	US-09-815-242-13581 Sequence 13581, A
45	116.5	13.2	246	9	US-09-738-626-5141 Sequence 5141, Ap

ALIGNMENTS

RESULT 1
US-09-860-846-33
; Sequence 33, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 3739
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-33

Query Match 62.0%; Score 547; DB 9; Length 3739;
Best Local Similarity 62.8%; Pred. No. 5-5e-40;

Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;

QY 1 GSVLVTGCTGVGAARHLCVCGVRLDILVSRGPDPAEGRLAEIALGAERYIVA 60
DB 3277 GSVLVTGCTGVGAARHLCVCGVRLDILVSRGPDPAEGRLAEIALGAERYIVA 3336
QY 61 CQVGRREVRRLDEVPACCPVTVHAAGVDDATISLTPERLCTVPAAVDALIID 120
DB 3337 CQVGRREVRRLDEVPACCPVTVHAAGVDDATISLTPERLCTVPAAVDALIID 3396
QY 121 ELTR-ANGELSAFVLFSSAAGILGAGOGNYAANAADALAYRRRAAGLPGVSLAMGLW 178
DB 3397 ELTR-ANGELSAFVLFSSAAGILGAGOGNYAANAADALAYRRRAAGLPGVSLAMGLW 3456
RESULT 2
US-09-988-384B-33

[illegible]

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FaestSeq for Windows Version 3.0
 SEQ ID NO 33
 LENGTH: 3739
 TYPE: PRF
 ORGANISM: Streptomyces venezuelae
 US-09-861-289-33

Query Match 62.0%; Score 547; DB 10; Length 3739;
 Best Local Similarity 62.8%; Pred. No. 5.5e-40;
 Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;

QY 1 GSVLTGCTGTVGAAGVAGVCGVRLDILVSRGPDPAAGAGLPAELALGAEVRIYA 60
 DB 3277 GTVLLTGCTGAGLGGIVARHVVGEMGVRRLLVSRGTDAFGAGELVHELEALGADVSVAA 3336
 QY 61 CDVGERREVRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLLD 120
 DB 3337 CDVADREALTAVLDSIPAEHPPLTAVVHTAGVLSDDGLPSMTAEDEVHVLRPKYDAAFLLD 3396
 QY 121 ELTR--GMELSAFVLFSSAAGILGSAGGQNYAANAALDALAYRRRAAGLPGVSLAMGLM 178
 DB 3397 ELTSTPGVDLAAFVWFSSAAAVFGAGGAGAYAAANATLDALAMRRRTAGLPALSLGWMGLM 3456

RESULT 6
 US-09-860-846-6
 Sequence 6, Application US/09860846
 Patent No. US20020164742A1
 GENERAL INFORMATION:
 APPLICANT: Sherman, D.H.
 APPLICANT: Liu, H.
 APPLICANT: Xue, Y.
 APPLICANT: Zhao, L.
 TITLE OF INVENTION: DNA encoding methymycin and pikromycin
 FILE REFERENCE: 600.438US1
 CURRENT APPLICATION NUMBER: US/09/860,846
 CURRENT FILING DATE: 2001-05-18
 PRIOR APPLICATION NUMBER: 09/105,537
 PRIOR FILING DATE: 1998-06-26
 NUMBER OF SEQ ID NOS: 43
 SOFTWARE: FaestSeq for Windows Version 3.0
 SEQ ID NO 6
 LENGTH: 11877
 TYPE: PRF
 ORGANISM: Streptomyces venezuelae
 US-09-860-846-6

Query Match 62.0%; Score 547; DB 9; Length 11877;
 Best Local Similarity 62.8%; Pred. No. 2.2e-39;
 Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;

QY 1 GSVLTGCTGTVGAAGVAGVCGVRLDILVSRGPDPAAGAGLPAELALGAEVRIYA 60
 DB 8226 GTVLLTGCTGAGLGGIVARHVVGEMGVRRLLVSRGTDAFGAGELVHELEALGADVSVAA 8285
 QY 61 CDVGERREVRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLLD 120
 DB 8286 CDVADREALTAVLDSIPAEHPPLTAVVHTAGVLSDDGLPSMTAEDEVHVLRPKYDAAFLLD 8345
 QY 121 ELTR--GMELSAFVLFSSAAGILGSAGGQNYAANAALDALAYRRRAAGLPGVSLAMGLM 178
 DB 8346 ELTSTPGVDLAAFVWFSSAAAVFGAGGAGAYAAANATLDALAMRRRTAGLPALSLGWMGLM 8405

RESULT 7
 US-09-836-821-6
 Sequence 6, Application US/09836821
 Patent No. US20030087405A1
 GENERAL INFORMATION:
 APPLICANT: Sherman, D.H.
 APPLICANT: Liu, H.
 APPLICANT: Xue, Y.
 APPLICANT: Zhao, L.

TITLE OF INVENTION: DNA encoding methymycin and pikromycin
 FILE REFERENCE: 600.438US1
 CURRENT APPLICATION NUMBER: US/09/836,821
 CURRENT FILING DATE: 2001-04-17
 PRIOR APPLICATION NUMBER: 09/105,537
 PRIOR FILING DATE: 1998-06-26
 NUMBER OF SEQ ID NOS: 43
 SOFTWARE: FaestSeq for Windows Version 3.0
 SEQ ID NO 6
 LENGTH: 11877
 TYPE: PRF
 ORGANISM: Streptomyces venezuelae
 US-09-836-821-6

Query Match 62.0%; Score 547; DB 9; Length 11877;
 Best Local Similarity 62.8%; Pred. No. 2.2e-39;
 Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;

QY 1 GSVLTGCTGTVGAAGVAGVCGVRLDILVSRGPDPAAGAGLPAELALGAEVRIYA 60
 DB 8226 GTVLLTGCTGAGLGGIVARHVVGEMGVRRLLVSRGTDAFGAGELVHELEALGADVSVAA 8285
 QY 61 CDVGERREVRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLLD 120
 DB 8286 CDVADREALTAVLDSIPAEHPPLTAVVHTAGVLSDDGLPSMTAEDEVHVLRPKYDAAFLLD 8345
 QY 121 ELTR--GMELSAFVLFSSAAGILGSAGGQNYAANAALDALAYRRRAAGLPGVSLAMGLM 178
 DB 8346 ELTSTPGVDLAAFVWFSSAAAVFGAGGAGAYAAANATLDALAMRRRTAGLPALSLGWMGLM 8405

RESULT 8
 US-09-861-289-6
 Sequence 6, Application US/09861289
 Patent No. US20020110897A1
 GENERAL INFORMATION:
 APPLICANT: Sherman, D.H.
 APPLICANT: Liu, H.
 APPLICANT: Xue, Y.
 APPLICANT: Zhao, L.
 TITLE OF INVENTION: DNA encoding methymycin and pikromycin
 FILE REFERENCE: 600.438US1
 CURRENT APPLICATION NUMBER: US/09/861,289
 CURRENT FILING DATE: 2001-05-18
 PRIOR APPLICATION NUMBER: 09/105,537
 PRIOR FILING DATE: 1998-06-26
 NUMBER OF SEQ ID NOS: 43
 SOFTWARE: FaestSeq for Windows Version 3.0
 SEQ ID NO 6
 LENGTH: 11877
 TYPE: PRF
 ORGANISM: Streptomyces venezuelae
 US-09-861-289-6

Query Match 62.0%; Score 547; DB 10; Length 11877;
 Best Local Similarity 62.8%; Pred. No. 2.2e-39;
 Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;

QY 1 GSVLTGCTGTVGAAGVAGVCGVRLDILVSRGPDPAAGAGLPAELALGAEVRIYA 60
 DB 8226 GTVLLTGCTGAGLGGIVARHVVGEMGVRRLLVSRGTDAFGAGELVHELEALGADVSVAA 8285
 QY 61 CDVGERREVRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLLD 120
 DB 8286 CDVADREALTAVLDSIPAEHPPLTAVVHTAGVLSDDGLPSMTAEDEVHVLRPKYDAAFLLD 8345
 QY 121 ELTR--GMELSAFVLFSSAAGILGSAGGQNYAANAALDALAYRRRAAGLPGVSLAMGLM 178
 DB 8346 ELTSTPGVDLAAFVWFSSAAAVFGAGGAGAYAAANATLDALAMRRRTAGLPALSLGWMGLM 8405

RESULT 9
 US-09-988-384B-6

```
; Sequence 6, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 6
; LENGTH: 12199
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-988-384B-6
```

```
Query Match          62.0%; Score 547; DB 9; Length 12199;
Best Local Similarity 62.8%; Pred. No. 2,3e-39;
Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;

QY 1 GSVLVTGCTGTVGAARHLAGVCGVRDLTVSRGPDAPGAEGLRAELAAAGAEVRIYA 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8548 GTVLTGCTGCTAGTGAARHVGEMGVRLLVSRGTDAPGAEGLRAELAAAGADVSVA 8607
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 CDVGERREVRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKYDALLD 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8608 CDVADREALTAVALDIPAEHPLTAVVHTAGVLSDTLPSTADVEHVLRPKYDAFLD 8667
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 ELTR--GMELSAFVLFSSAAGILGSAAGCNVAAAANALDALYRRRAAGLPGVSLAWGLM 178
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8668 ELTSTPAYDLAAVFWFSSAAAVFGAGGAGVAAANATLDALAMRRRAAGLPALSLGWMGLM 8727
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 10
US-09-793-708-1
; Sequence 1, Application US/09793708
; Publication No. US20030104597A1
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002121
; CURRENT APPLICATION NUMBER: US/09/793,708
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 09/657,440
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: US 09/141,908
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: US 09/073,538
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: US 08/846,247
; PRIOR FILING DATE: 1997-04-30
; PRIOR APPLICATION NUMBER: US 60/134,990
; PRIOR FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4551
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-793-708-1
```

```
Query Match          61.9%; Score 546; DB 9; Length 4551;
Best Local Similarity 62.8%; Pred. No. 8,5e-40;
Matches 113; Conservative 20; Mismatches 45; Indels 2; Gaps 1;

QY 1 GSVLVTGCTGTVGAARHLAGVCGVRDLTVSRGPDAPGAEGLRAELAAAGAEVRIYA 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4038 GTVLTGCTGCTAGTGAARHVGEMGVRLLVSRGTDAPGAEGLRAELAAAGADVSVA 4097
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 CDVGERREVRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKYDALLD 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4098 CDVADREALTAVALDIPAEHPLTAVVHTAGVLSDTLPSTADVEHVLRPKYDAFLD 4157
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 ELTR--GMELSAFVLFSSAAGILGSAAGCNVAAAANALDALYRRRAAGLPGVSLAWGLM 178
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4158 ELTSTPAYDLAAVFWFSSAAAVFGAGGAGVAAANATLDALAMRRRAAGLPALSLGWMGLM 4217
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 11
US-09-860-846-31
; Sequence 31, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-860-846-31
```

```
Query Match          61.9%; Score 546; DB 9; Length 4613;
Best Local Similarity 62.8%; Pred. No. 8,7e-40;
Matches 113; Conservative 20; Mismatches 45; Indels 2; Gaps 1;

QY 1 GSVLVTGCTGTVGAARHLAGVCGVRDLTVSRGPDAPGAEGLRAELAAAGAEVRIYA 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4100 GTVLTGCTGCTAGTGAARHVGEMGVRLLVSRGTDAPGAEGLRAELAAAGADVSVA 4159
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 CDVGERREVRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKYDALLD 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4160 CDVADREALTAVALDIPAEHPLTAVVHTAGVLSDTLPSTADVEHVLRPKYDAFLD 4219
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 ELTR--GMELSAFVLFSSAAGILGSAAGCNVAAAANALDALYRRRAAGLPGVSLAWGLM 178
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4220 ELTSTPAYDLAAVFWFSSAAAVFGAGGAGVAAANATLDALAMRRRAAGLPALSLGWMGLM 4279
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 12
US-09-988-384B-31
; Sequence 31, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
```


Job time : 6.54613 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 17, 2003, 12:54:32 ; Search time 3.2324 Seconds
(without alignments)
1618.821 Million cell updates/sec

Title: US-09-914-286-3_COPY_3537_3714
Perfect score: 882
Sequence: 1 GSVLTGCGTGVGAIVARHL.....LAYRRRAAGLPGVSLAWGLW 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfillseq.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	568	64.4	3724	2	US-08-804-227C-10
2	568	64.4	3724	2	US-08-804-198-4
3	547	62.0	3729	3	US-09-320-878-2
4	547	62.0	3729	4	US-09-105-537-33
5	547	62.0	11877	4	US-09-105-537-6
6	546	61.9	4551	3	US-09-320-878-1
7	546	61.9	4613	4	US-09-105-537-31
8	534	60.5	4472	2	US-08-804-227C-2
9	530	59.0	5215	4	US-09-105-537-2
10	520	59.0	5588	4	US-09-036-987A-6
11	520	59.0	5588	4	US-09-370-700-6
12	518	58.7	1864	2	US-08-804-227C-3
13	512.5	58.1	6095	4	US-09-144-085-2
14	509.5	57.8	5087	4	US-09-144-085-1
15	504	57.1	4928	4	US-09-036-987A-5
16	504	57.1	4928	4	US-09-370-700-5
17	500	56.7	4550	2	US-08-804-227C-14
18	500	56.7	4550	2	US-08-804-227C-8
19	500	56.7	4550	2	US-08-804-198-2
20	483.5	54.8	1841	2	US-08-804-227C-6
21	477	54.6	3729	2	US-08-804-227C-4
22	477	54.1	1996	2	US-08-804-227C-9
23	477	53.9	1996	2	US-08-804-198-3
24	475	53.9	2152	4	US-09-036-987A-3
25	475	53.9	2152	4	US-09-370-700-3
26	475	53.9	3567	2	US-07-642-734C-4
27	475	53.9	3567	3	US-08-439-009A-4

28	463.5	52.6	3170	4	US-09-036-987A-4	Sequence 4, App11
29	463.5	52.6	3170	4	US-09-370-700-4	Sequence 4, App11
30	449.5	51.0	3816	4	US-09-428-517-3	Sequence 3, App11
31	440.5	49.9	7257	3	US-09-335-409-5	Sequence 5, App11
32	440.5	49.9	7257	4	US-09-568-102-5	Sequence 5, App11
33	440.5	49.9	7257	4	US-09-567-969-5	Sequence 5, App11
34	440.5	49.9	7257	4	US-09-568-480-5	Sequence 5, App11
35	440.5	49.9	7257	4	US-09-568-486-5	Sequence 5, App11
36	440.5	49.9	7257	4	US-09-568-472-5	Sequence 5, App11
37	440.5	49.9	7257	4	US-09-567-899-5	Sequence 5, App11
38	438	49.7	3170	2	US-07-642-734C-5	Sequence 5, App11
39	438	49.7	3170	3	US-08-439-009A-5	Sequence 5, App11
40	431	48.9	159	3	US-09-010-809-1	Sequence 1, App11
41	428.5	48.6	2595	4	US-09-036-987A-2	Sequence 2, App11
42	428.5	48.6	2595	4	US-09-370-700-2	Sequence 2, App11
43	428	48.5	4150	4	US-09-428-517-2	Sequence 2, App11
44	426.5	48.4	3491	2	US-07-642-734C-2	Sequence 2, App11
45	426.5	48.4	3491	3	US-08-439-009A-2	Sequence 2, App11

ALIGNMENTS

RESULT 1
US-08-804-227C-10
; Sequence 10, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Khusose, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3724 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-804-227C-10
Query Match 64.4%; Score 568; DB 2; Length 3724;
Best Local Similarity 65.2%; Pred. No. 66-54;
Matches 116; Conservative 19; Mismatches 43; Indels 0; Gaps 0;
QY 1 GSVLTGCGTGVGAIVARHLGVRRGPDAGAGAPRATAGAVRIVA 60
DB 3268 GSVLTGCGTGVGAIVARHLGVRRGPDAGAGAPRATAGAVRIVA 3327
QY 61 CDVGRREVVRLLGVGPAGCPVTGVVHAAGVLDATIASLTPTERTVFAKVDALLLD 120

DB 3328 CDVSDADAVRGLAGIPADHPLAIVHSGVLDGVLPGILTPRRMRVLRPKVEAAVHLD 3387
QY 121 ELTRGMEISAFVLFSSAGILGSAGOGNYAANAALDALAYRRRAGLPGVSLAMGLW 178
DB 3388 ELTRDLDLSAFVLFSSAGILGSAGOGNYAANAATLDALAARRRSLGLPSVSLAMGLW 3445

RESULT 2

US-08-804-198-4
; Sequence 4, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rostock, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLOIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL, 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3724 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-804-198-4

Query Match 64.4%; Score 568; DB 2; Length 3724;
Best Local Similarity 65.2%; Pred. No. 6e-54;
Matches 116; Conservative 19; Mismatches 43; Indels 0; Gaps 0;
QY 1 GSVLVGTGTVGAAGAAVAAHLAGVCGVRDILLVSRGPDAPGAGLRAELAAAGAEVRIYA 60
DB 3268 GTVLVTGCTGALGRIARHLASRHGVRLLIARRGPDGGAELVADLAALGASATVYA 3327
QY 61 CDVGERREVRLLLEGVPACPLTGVVHAAGVLDATIAALTPERLGTVAAPKVDALLLD 120
DB 3328 CDVSDADAVRGLAGIPADHPLAIVHSGVLDGVLPGILTPRRMRVLRPKVEAAVHLD 3387
QY 121 ELTRGMEISAFVLFSSAGILGSAGOGNYAANAALDALAYRRRAGLPGVSLAMGLW 178
DB 3388 ELTRDLDLSAFVLFSSAGILGSAGOGNYAANAATLDALAARRRSLGLPSVSLAMGLW 3445

RESULT 3
US-09-320-878-2
; Sequence 2, Application US/09320878A
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary

; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: McDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARONOLOIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; EARLIER FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 3739
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-320-878-2

Query Match 62.0%; Score 547; DB 3; Length 3739;
Best Local Similarity 62.8%; Pred. No. 1.4e-51;
Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;

QY 1 GSVLVGTGTVGAAGAAVAAHLAGVCGVRDILLVSRGPDAPGAGLRAELAAAGAEVRIYA 60
DB 3277 GTVLVTGCTGALGRIARHLASRHGVRLLIARRGPDGGAELVADLAALGASATVYA 3336
QY 61 CDVGERREVRLLLEGVPACPLTGVVHAAGVLDATIAALTPERLGTVAAPKVDALLLD 120
DB 3337 CDVADREALTVLDSIPAEHPLTAIVHAGVLDGTLPSMTAEVDEHVLRPKVDAAFLLD 3396
QY 121 ELTR--GHELSAFVLFSSAGILGSAGOGNYAANAALDALAYRRRAGLPGVSLAMGLW 178
DB 3397 ELTSTPGVDLAFVWFSSAAVFGAGGAYAANAATLDALAMRRRTAGLPALSLGWGLW 3456

RESULT 4
US-09-105-537-33
; Sequence 33, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 3739
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-33

Query Match 62.0%; Score 547; DB 4; Length 3739;
Best Local Similarity 62.8%; Pred. No. 1.4e-51;
Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;

QY 1 GSVLVGTGTVGAAGAAVAAHLAGVCGVRDILLVSRGPDAPGAGLRAELAAAGAEVRIYA 60
DB 3277 GTVLVTGCTGALGRIARHLASRHGVRLLIARRGPDGGAELVADLAALGASATVYA 3336


```

Query Match      61.9%; Score 546; DB 4; Length 4613;
Best Local Similarity 62.8%; Pred. No. 2,4e-51;
Matches 113; Conservative 20; Mismatches 45; Indels 2; Gaps 1;

Qy      1 GSVLVTGGTGLGAARVRLHAGVGVVDLLVSRHGDPAGEGLRAELMALGAERIVA 60
       |:|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      4100 GTVLTTGTGTLGGSLVARHVVGKGVRRLLLVSRGTDACGDELVHELRLGADVVA 4159

Qy      61 CDNGEREVRVRLLEGVPAGCPLTGVVHAAGVLDATTASLTPERLTGVFAAKVDALLD 120
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      4160 CDVADREALTAVALDAIPEAHPLTAVHTAGLSPTSPMTTEVEVHLAPKVDAAFLLD 4219

Qy      121 ELNR--GMELSAFLPFSAGGIISGAGQGYAAANALDLAAYRRRAAGHPGVSLAWGLW 178
       |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      4220 ELISTPAYDALAIFMBSAAVAFQGAGOGAYAAANNATLLDALAMRRRAAGLPALSLGWGLW 4279

RESULT 8
US-08-804-227C-2
Sequence 2, Application US/08804227C
Patent No. 5876591
GENERAL INFORMATION:
Applicant: Dehoff, Bradley S.
Applicant: Kuhsees, Stuart A.
Applicant: Rosteck, Paul R., Jr.
Applicant: Sutton, Kimberly L.
```

```

/ TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
/ NUMBER OF SEQUENCES: 15
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: THOMAS G. PLANT 1501
/ STREET: LULUY CORPORATE CENTER
/ CITY: INDIANAPOLIS
/ STATE: IN
/ COUNTRY: USA
/ ZIP: 46285
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: ASCII(DOS) Text only
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/804.227C
/ FILING DATE: February 21, 1997
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Plant, Thomas, G.
/ REGISTRATION NUMBER: 35,784
/ REFERENCE/DOCKET NUMBER: X-8231
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 317-276-2459
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4472 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/
US-08-804-227C-2

```

Query Match 60.5%; Score 534; DB 2; Length 4472;
Best Local Similarity 64.0%; Pred. No. 5e-50;

Matches 114; Conservative 12; Mismatches 52; Indels 0; Gaps 0;

```

QY 1 GSVLVGTGTGVLGAARHLAGVCGVRDLLVSRGPDAPGAGLRRAELALGAERYIVA 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4017 GTVLVTGATGTGLRLRLARHLAGVGRHLLLSRGGRAAGADELAELAEPCFMA 4076
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLD 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4077 CDADRESLARLVAEPVADPLGVHAAVLDGTLDTLTPRIGTVMPKADALNLH 4136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 ELTRGHELSAPFLFSSAGILSGAGCGNVAANAALDALAYRRRAAGLPGVSLANGLM 178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4137 ELTRISPLSVFAVFSQAAGILGRPGQANTAAATFDALQHRRAAGLPAVSLANGLM 4194
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 9
US-09-105-537-2
/ Sequence 2, Application US/09105537A
/ Patent No. 6265202
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, D.H.
/ APPLICANT: Liu, H.
/ APPLICANT: Xue, Y.
/ APPLICANT: Zhao, L.
/ TITLE OF INVENTION: DNA encoding methymycin and pikromycin
/ FILE REFERENCE: 600.4380U1
/ CURRENT APPLICATION NUMBER: US/09/105.537A
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2
/ LENGTH: 5215
/ TYPE: PRT
/ ORGANISM: Streptomyces venezuelae
/
US-09-105-537-2

```

Query Match 60.1%; Score 530; DB 4; Length 5215;
Best Local Similarity 62.9%; Pred. No. 1.7e-49;
Matches 112; Conservative 14; Mismatches 52; Indels 0; Gaps 0;

```

QY 1 GSVLVGTGTGVLGAARHLAGVCGVRDLLVSRGPDAPGAGLRRAELALGAERYIVA 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2440 GTVLVTGATGTGALLARLVERHGRRLILVSRGADAPGADGEDLAGLGAERYAFNA 2499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLD 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2500 ADADRESLARLVAEPVADPLGVHAAVLDGTLDTLTPRIGTVMPKADALNLH 2559
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 ELTRGHELSAPFLFSSAGILSGAGCGNVAANAALDALAYRRRAAGLPGVSLANGLM 178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2560 ELTKDLRLDAFVLFSVSGIVGTAGQANYAANTGIDALAHRAATGATSLANGLM 2617
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Query Match 59.0%; Score 520; DB 4; Length 5588;
Best Local Similarity 62.9%; Pred. No. 2.5e-48;

Matches 112; Conservative 16; Mismatches 50; Indels 0; Gaps 0;

```

QY 1 GSVLVGTGTGVLGAARHLAGVCGVRDLLVSRGPDAPGAGLRRAELALGAERYIVA 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4933 GTVLVTGATGTGALLARLVERHGRRLILVSRGADAPGADGEDLAGLGAERYAFNA 4992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLD 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4993 CDADRESLARLVAEPVADPLGVHAAVLDGTLDTLTPRIGTVMPKADALNLH 5052
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

US-09-036-987A-6

Sequence 6, Application US/09036987A
Patent No. 6143526
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patli J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionaville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036.987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 5588 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-6

```
Oy      121 ELTRGHELSAFLVFPSSAGILSGAGCGNTAAANPALDALAYRRRAAGLPGVSLAWGIW 178
        ||| : ||||| | : ||||| | : ||||| | : ||||| | : ||||| | : ||||| |
Db      5053 ELTRREGLSAPLVFPSSVAGLPFGAGQGSNTAAAGNAFLDALHCRQAQGLPALSASGIW 5110
```

RESULT 11
US-09-370

```

US-09-370-700-6
Sequence 6, Application US/09370700
Patent No. 6274350
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patti J
APPLICANT: Turner, Jan R
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIVI
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 5588
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-6

```

Query Match 59.0%; Score 520; DB 4; Length 5588;
Best Local Similarity 62.9%; Pred. No. 2.5e+48;
Matches 112; Conservative 16; Mismatches 50; Indels 0; Gaps 0;

Qy 1 GSVLYTGTGYLGAANVAHRLHVLGVCVRDLIVSRGPAQPAQEGRIAEIAALAGAIVRIYA 60
::: :::
Db 4993 GTVLYTGTGTGGALNARHLIVSAHGVRLHLILASRRGLAAGDAELVALEEQAGDAVAVYA 49992

Qy 61 CDVGERRRVRLLEGVPAQCEPLTGVNNAAGVLDPAITISLPERGCTPAKRVADALLDD 120
::: :::
Db 4993 CDSADRRSLARIVASVPAENELRVVYHAAQVLDGVLMSRERIDAVLRPKYDAWKLH 5052

Qy 121 ELTGMELISAFVLFPSSAGILGSAQCSYAAANAALDALAYRRRAGLPGVSLANGWL 178
::: :::
Db 5053 ELTRELGLISAFVLFPSSVAGLFCGAAQCSYAAAGNAETDLALHQRQAQGLPALSLSAGWL 510
::: :::

RESULT 12

US-08-804-227C-3
 1 Sequence 3, Application US/08804227C
 2 Patent No. 5876591
 3
 4 GENERAL INFORMATION:
 5 APPLICANT: Dehoff, Bradley S.
 6 APPLICANT: Kuhseos, Stuart A.
 7 APPLICANT: Roeseck, Paul R., Jr.
 8 APPLICANT: Sutton, Kimberly L.
 9 TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
 10
 11 NUMBER OF SEQUENCES: 15
 12
 13 CORRESPONDENCE ADDRESS:
 14 ADDRESSEE: THOMAS G. PLANT 1501
 15 STREET: LILLY CORPORATE CENTER
 16 CITY: INDIANAPOLIS
 17
 18 STATE: IN
 19
 20 COUNTRY: USA
 21
 22 ZIP: 46285
 23
 24 COMPUTER READABLE FORM:
 25 MEDIUM TYPE: Floppy disk
 26 COMPUTER: IBM Compatible
 27 OPERATING SYSTEM: MS-DOS
 28 SOFTWARE: ASCII(DOS) Text only
 29
 30 CURRENT APPLICATION DATA:
 31 APPLICATION NUMBER: US/08/804,227C

? FILING DATE: February 21, 1997
 ? CLASSIFICATION: 435
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Plant, Thomas, G.
 ? REGISTRATION NUMBER: 35,784
 ? REFERENCE/DOCKET NUMBER: X-823
 ? TELECOMMUNICATION INFORMATION:
 ?

Query Match	58.7%	Score 518	DB 2	Length 1864
Best Local Similarity	60.7%	Pred. No. 8.7e-49		
Matches 108	Conservative 16	Mismatches 54	Indels 0	Gaps 0

Oy	1	GSVLVTTGGTGVLGAAVVRHVLHAGVCGBDILLVSRRGDAPABEGLRAELALGAEVIVA	60
		::: ::: ::: ::: ::: ::: ::: :::	
Dd	1414	GTIVITIGTGSSLGEAARHHLLVTRHGARRLLITSRSGQAPGAABLVAELALGHADYVA	1473
		::: ::: ::: ::: ::: ::: ::: :::	
Oy	61	CDVGERREVRVLTLEGVPACPLTGVVRAAGVLDPATIASLPBELGTVPFAKVDNALILD	120
		::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	1474	CDVADRALLRALDLRVPRGHEPLTVLITAGVLDGCVLTAQTPOQLAAVLRPKDAVRNLH	1533
		::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Oy	121	ELTRGMELSAFVLPSSAAGIISGAGCGNVAANAALDALAVRRRAAGPGVSLANGIM	178
		::: ::: ::: ::: ::: ::: ::: :::	
Dd	1534	ELTGGHALSAILTVSSAAGVLSGAGSGCYAANAAYLUSPAVMKRSRGLPAYSILDMGPV	1591

RESULT 13
US-09-144

Sequence 2, Application US/09144085
Patent No. 6280999
GENERAL INFORMATION:
APPLICANT: Gustafsson, Claes
APPLICANT: Betlach, Mary C.
APPLICANT: Ashley, Gary
APPLICANT: Julien, Bryan
APPLICANT: Ziemann, Rainer
TITLE OF INVENTION: SORANGLUM POLYKETIDE SYNTHASES AND ENCODING DNA
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: 30062-20020.20
CURRENT APPLICATION NUMBER: US/09/144,085
CURRENT FILING DATE: 1998-08-31
EARLIER APPLICATION NUMBER: 09/010,809
EARLIER FILING DATE: 1998-01-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 6095
TYPE: PR1
ORGANISM: Sorangium cellulosum
US-09-144-085-2

Query Match	Score	DB	Length
Best Local Similarity	58.1%;	4;	6095;
	60.9%;	Pred. No. 2e-47;	

[illegible]

RESULT 14

US-09-144-085-1

Sequence 1, Application US/09144085
Patent No. 6280999
GENERAL INFORMATION:
APPLICANT: Gustafsson, Claes
APPLICANT: Gustafsson, Mary C.
APPLICANT: Ashley, Gary
APPLICANT: Julien, Bryan
APPLICANT: Ziemann, Rainer
TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
FILE REFERENCE: 30062-20020.20
CURRENT APPLICATION NUMBER: US/09/144,085
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 5087
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-144-085-1

Query Match 57.8%; Score 509.5; DB 4; Length 5087;

Best Local Similarity 61.5%; Pred. No. 3.3e-47; Matches 110; Conservative 13; Mismatches 55; Indels 1; Gaps 1;

QY 1 GSVLTGTTGTVGAAYARHLAGVCGVRDLLVSRGPDAPGAEGLRAELALGAE-VRIY 59
DB 4642 GTVLVTGTTGELQAAVAHLYRAHGVRIHVLTSRRGLEAGAEIVESLAEIGAEVYVA 4701
QY 60 ACVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLI 119
DB 4702 ACVSRKREEVARVLGIDARPLSAVHLALGALDDGVLGQTAERLSRVLAPEVDALHL 4761
QY 120 DELTRGMELSAFVLFSSAGILGSAAGVAAANALDMLAYRRRAAGLPGVSLANGWL 178
DB 4762 HELTRLELDVAFVLFSSVAGTFTGAGOSNTAAANTFLDALAHRRGGLAATSLANGWL 4820

RESULT 15

US-09-036-987A-5

Sequence 5, Application US/09036987A
Patent No. 6143526
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/036,987A

FILING DATE: 09-MAR-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stuart, Donald R

REGISTRATION NUMBER: 28,479

REFERENCE/DOCKET NUMBER: 50,608

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317)337-4816

TELEFAX: (317)337-4847

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 4928 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-036-987A-5

Query Match 57.1%; Score 504; DB 4; Length 4928;

Best Local Similarity 59.0%; Pred. No. 1.3e-46; Matches 105; Conservative 19; Mismatches 54; Indels 0; Gaps 0;

QY 1 GSVLTGTTGTVGAAYARHLAGVCGVRDLLVSRGPDAPGAEGLRAELALGAEVRIYA 60
DB 1421 GTVLITGTTGALGVVARRHVEHGRRLVLAGRRGMNAPGVHELVDLARAGAVEVVA 1480
QY 61 CVVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDALLI 120
DB 1481 CVVADRTDEHVLALIPVDWPLRGIVHTRGVADGYTSLAADYGVTFAPKVTGAMHLH 1540
QY 121 ELTRGMELSAFVLFSSAGILGSAAGVAAANALDMLAYRRRAAGLPGVSLANGWL 178
DB 1541 ELTRDLDLSEFFVLFSSFSIGAGAGVAAANNTFLDALARRRARGLPGLSLANGWL 1598

Search completed: June 17, 2003, 13:09:31
Job time: 4.23524 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 17, 2003, 12:50:57 ; Search time 8.71975 Seconds
(without alignments)
2720.102 Million cell updates/sec

Title: US-09-914-286-3_COPY_3537_3714
Perfect score: 882
Sequence: 1 GSVLVTGCGTGLGAVALRH.....LAYRRRAAGLPGVSLAWGLM 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A.GeneSeq_101002.*
2: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.*
3: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.*
4: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.*
5: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.*
6: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.*
7: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.*
8: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.*
9: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.*
10: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.*
11: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.*
12: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.*
13: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.*
14: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.*
15: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.*
16: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.*
17: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.*
18: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.*
19: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.*
20: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.*
21: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	882	100.0	3972	21	AA823749
2	882	100.0	3972	22	AA65264
3	882	100.0	3972	22	AA65268
4	882	100.0	5532	21	AA823752
5	882	100.0	5532	22	AA65267
6	616	69.8	6239	21	AA823750
7	616	69.8	6239	22	AA65265
8	568	64.4	3724	18	AAW23718
9	568	64.4	3724	18	AAW22608
10	554	62.8	2066	22	AA810146

11	549	62.2	11096	22	AA810129	Streptomyces nours
12	548.5	62.2	9477	22	AA810144	Streptomyces nours
13	547	62.0	3739	21	AA818638	Amino acid sequenc
14	547	62.0	3739	21	AAV77193	S. venezuelae macr
15	547	62.0	3739	21	AAV77201	S. venezuelae pik
16	547	62.0	3739	21	AAV77202	Natobionide synth
17	547	62.0	12199	21	AAV77180	S. venezuelae pik
18	546	61.9	4551	21	AA818637	Amino acid sequenc
19	546	61.9	4551	21	AAV67201	Natobionide synth
20	546	61.9	4613	21	AAV77192	S. venezuelae macr
21	546	61.9	4613	21	AAV77200	S. venezuelae pik
22	546	61.9	5435	22	AA810145	Streptomyces nours
23	537	60.9	899	22	AA810153	Protein encoded by
24	534	60.5	4472	18	AAW22601	Tylosone synthase
25	530	60.1	4630	21	AAV77177	S. venezuelae vep
26	523	59.3	4630	18	AAW19629	Streptomyces venez
27	520	59.0	4572	19	AAW52845	A. mediterranei ri
28	520	59.0	5588	20	AAV39301	Sphs a polyketide
29	520	59.0	5588	22	AAW70969	S. spinosa protein
30	518	58.7	1864	18	AAW22602	Tylosone synthase
31	517	58.6	7068	22	AA810142	Streptomyces nours
32	512.5	58.1	6095	22	AAU10701	Amino acid sequenc
33	511	57.9	6797	22	AA831558	Platencin biosynth
34	509.5	57.8	5087	22	AAU10700	Amino acid sequenc
35	504	57.1	4924	22	AAW70968	S. spinosa protein
36	504	57.1	1891	20	AAV39300	Sphd a polyketide
37	503.5	57.1	1891	19	AAW52844	Amvicolatopsis med
38	503.5	57.1	3413	19	AAW52849	A. mediterranei ri
39	500	56.7	4545	18	AAW22611	Hydrid 8mg/tylg O
40	500	56.7	4550	18	AAW23716	Platenolide synth
41	500	56.7	4550	18	AAW22606	Platenolide synth
42	495.5	56.2	1688	19	AAW52848	A. mediterranei ri
43	492	55.8	5069	19	AAW52846	A. mediterranei ri
44	481	55.7	4881	21	AA823751	S. avermectilis ave
45	491	55.7	4881	22	AA65266	Streptomyces averm

ALIGNMENTS

RESULT 1
AAB23749 standard; Protein; 3972 AA.
ID: AAB23749
XX
AC AAB23749;
XX
DT 10-JAN-2001 (first entry)
XX
DE S. avermectilis avermectin aglycon synthase protein SEQ ID NO:3.
XX
KW Streptomyces avermectilis; avermectin aglycon synthase; bioisynthetis;
KW multifunctional enzyme; polyketide; avermectin; veterinary drug;
KW agrochemical.
XX
OS Streptomyces avermectilis.
XX
PN WO200050605-A1.
XX
PD 31-AUG-2000.
XX
PF 23-FEB-2000; 2000WO-JP01041.
XX
PR 24-FEB-1999; 99JP-0046961.
XX
PA (KITA) KITASATO INST.
XX
PI Omura S, Ikeda H;
XX
DR WPI: 2000-565458/52.
XX
DR N-PSDB; AAA92301;
XX
PT Avermectin aglycone synthase DNA and proteins encoded by all or part of
it for the production of avermectin and its derivatives for drug and

PT agrochemical use -
XX
XX Claim 32; Page 203-223; 314pp. Japanese.
XX
CC The present sequence represents an avermectin aglycon synthase protein.
CC Also described are: (1) polypeptides encoded by all or part of the DNA;
CC (2) expression vectors containing the DNA; (3) host cells transformed by
CC the vectors; (4) preparation of the polypeptides by culture of the
CC transformants; (5) preparation of avermectin aglycon or its derivatives
CC by culture of transformed avermectin-producing microorganisms; and (6)
CC oligonucleotides of 5-60 bases in length containing sense or antisense
CC sequences from the avermectin aglycon synthase DNA. The enzymes are
CC useful for the production of modified forms of avermectin and of the
CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
CC and agrochemicals.
XX
SQ Sequence 3972 AA:

Query Match 100.0%; Score 882; DB 21; Length 3972;
Best Local Similarity 100.0%; Pred. No. 1.8e-76;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSVLVNGTGVGAARHLAGVGVDDLLVSRGPDAGAGRLAEALAGAEVRIYA 60
DB 3537 GSVLVNGTGVGAARHLAGVGVDDLLVSRGPDAGAGRLAEALAGAEVRIYA 3596

QY 61 CDVGERREVRRLLEGVPAGCPLTGVAAGVLDATIASLTPERLGTVPAAKVDAALLD 120
DB 3597 CDVGERREVRRLLEGVPAGCPLTGVAAGVLDATIASLTPERLGTVPAAKVDAALLD 3656

QY 121 ELTRGMELSAFVLPSSAAGILGSAGGNTAAANAALDALAYRRRAAGLPGVSLAMGLM 178
DB 3657 ELTRGMELSAFVLPSSAAGILGSAGGNTAAANAALDALAYRRRAAGLPGVSLAMGLM 3714

RESULT 2
AAG65264
ID AAG65264 standard; Protein; 3972 AA.
XX
XX AAG65264;
XX
DT 04-DEC-2001 (first entry)
XX
DE Streptomycetes avermilitilis protein SEQ ID NO: 4.
XX
KM Avermectin aglycone synthase; AAS; avermectin derivative;
XX drug production; veterinary drug; pesticide.
OS Streptomycetes avermilitilis.
XX
PN MO200162939-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001MO-JP01381.
XX
PR 24-FEB-2000; 2000JP-0047405.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PA (KITA) KITASATO INST.
XX
PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
PI WPI; 2001-582053/65.
DR N-PSDB; AAH79277.
XX
XX New modified avermectin aglycone synthase derived from Streptomycetes
PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
PT drugs and pesticides -
XX
PS Claim 4; Page 167-180; 257pp; Japanese.
XX
CC The present invention relates to the production of modified derivatives

CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
CC avermilitilis. The activity of an acyl carrier protein (ACP),
CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC suppressed. The process can be used in the production of drugs, veterinary
CC drugs and pesticides. The present sequence is an S. avermilitilis protein.
XX
SQ Sequence 3972 AA:

Query Match 100.0%; Score 882; DB 22; Length 3972;
Best Local Similarity 100.0%; Pred. No. 1.8e-76;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSVLVNGTGVGAARHLAGVGVDDLLVSRGPDAGAGRLAEALAGAEVRIYA 60
DB 3537 GSVLVNGTGVGAARHLAGVGVDDLLVSRGPDAGAGRLAEALAGAEVRIYA 3596

QY 61 CDVGERREVRRLLEGVPAGCPLTGVAAGVLDATIASLTPERLGTVPAAKVDAALLD 120
DB 3597 CDVGERREVRRLLEGVPAGCPLTGVAAGVLDATIASLTPERLGTVPAAKVDAALLD 3656

QY 121 ELTRGMELSAFVLPSSAAGILGSAGGNTAAANAALDALAYRRRAAGLPGVSLAMGLM 178
DB 3657 ELTRGMELSAFVLPSSAAGILGSAGGNTAAANAALDALAYRRRAAGLPGVSLAMGLM 3714

RESULT 3
AAG65268
ID AAG65268 standard; Protein; 3972 AA.
XX
XX AAG65268;
XX
DT 04-DEC-2001 (first entry)
XX
DE Streptomycetes avermilitilis protein derivative SEQ ID NO: 8.
XX
KM Avermectin aglycone synthase; AAS; avermectin derivative;
XX drug production; veterinary drug; pesticide.
OS Synthetic.
XX
PN MO200162939-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001MO-JP01381.
XX
PR 24-FEB-2000; 2000JP-0047405.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PA (KITA) KITASATO INST.
XX
PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
PI WPI; 2001-582053/65.
DR N-PSDB; AAH79279.
XX
XX New modified avermectin aglycone synthase derived from Streptomycetes
PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
PT drugs and pesticides -
XX
PS Claim 5; Page 235-248; 257pp; Japanese.
XX
XX The present invention relates to the production of modified derivatives
CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
CC avermilitilis. The activity of an acyl carrier protein (ACP),
CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC suppressed. The process can be used in the production of drugs, veterinary
CC drugs and pesticides. The present sequence is derived from an S.
CC avermilitilis protein.

XX Sequence 3972 AA;
 SQ Query Match 100.0%; Score 882; DB 22; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 1.8e-76;
 Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSVLVTGCTGVLGAARHLAGVCGVRDLVSRGPDAPGAGLRAELAAAGAEVRIYA 60
 DB 3537 GSVLVTGCTGVLGAARHLAGVCGVRDLVSRGPDAPGAGLRAELAAAGAEVRIYA 3596

QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVFAAKVDALLLD 120
 DB 3597 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVFAAKVDALLLD 3656

QY 121 ELTRGMELSAFVLPSSAAGILGSGGNYAAANAALDALYRRRAAGLPGVSLAMGLM 178
 DB 3657 ELTRGMELSAFVLPSSAAGILGSGGNYAAANAALDALYRRRAAGLPGVSLAMGLM 3714

RESULT 4
 AAB23752 standard; Protein; 5532 AA.
 ID AAB23752;
 AC AAB23752;
 DT 10-JAN-2001 (first entry)
 XX 8. avermectin aglycon synthase protein SEQ ID NO:6.
 DE Streptomyces avermectin aglycon synthase; bioassay;
 KM Streptomyces avermectin aglycon synthase; bioassay;
 KM Streptomyces avermectin aglycon synthase; bioassay;
 KM Streptomyces avermectin aglycon synthase; bioassay;
 OS Streptomyces avermectin aglycon synthase.
 XX MO200050605-A1.
 PN 31-AUG-2000.
 PD 23-FEB-2000; 2000MO-JP01041.
 XX 24-FEB-1999; 99JP-0046961.
 XX (KITA) KITASATO INST.
 PA Omura S, Ikeda H;
 PI WPI; 2000-565458/52.
 DR N-PSDB; AAA92302.
 XX Avermectin aglycon synthase DNA and proteins encoded by all or part of
 PT it for the production of avermectin and its derivatives for drug and
 PT agrochemical use -
 XX Claim 32; Page 278-305; 314pp; Japanese.

XX The present sequence represents an avermectin aglycon synthase protein.
 CC Also described are: (1) polypeptides encoded by all or part of the DNA;
 CC (2) expression vectors containing the DNA; (3) host cells transformed by
 CC the vectors; (4) preparation of the polypeptides by culture of the
 CC transformants; (5) preparation of avermectin aglycon or its derivatives
 CC by culture of transformed avermectin-producing microorganisms; and (6)
 CC oligonucleotides of 5-60 bases in length containing sense or antisense
 CC sequences from the avermectin aglycon synthase DNA. The enzymes are
 CC useful for the production of modified forms of avermectin and of the
 CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
 CC and agrochemicals.

XX Sequence 5532 AA;
 SQ Query Match 100.0%; Score 882; DB 21; Length 5532;
 Best Local Similarity 100.0%; Pred. No. 2.7e-76;

Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSVLVTGCTGVLGAARHLAGVCGVRDLVSRGPDAPGAGLRAELAAAGAEVRIYA 60
 DB 1488 GSVLVTGCTGVLGAARHLAGVCGVRDLVSRGPDAPGAGLRAELAAAGAEVRIYA 1547

QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVFAAKVDALLLD 120
 DB 1548 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVFAAKVDALLLD 1607

QY 121 ELTRGMELSAFVLPSSAAGILGSGGNYAAANAALDALYRRRAAGLPGVSLAMGLM 178
 DB 1608 ELTRGMELSAFVLPSSAAGILGSGGNYAAANAALDALYRRRAAGLPGVSLAMGLM 1665

RESULT 5
 AAG65267 standard; Protein; 5532 AA.
 ID AAG65267;
 AC AAG65267;
 DT 04-DEC-2001 (first entry)
 XX Streptomyces avermectin aglycon synthase protein SEQ ID NO: 7.
 DE Streptomyces avermectin aglycon synthase; avermectin derivative;
 KM Avermectin aglycon synthase; avermectin derivative;
 KM Avermectin aglycon synthase; avermectin derivative;
 OS Streptomyces avermectin aglycon synthase.
 XX MO200162939-A1.
 PN 30-AUG-2001.
 PD 23-FEB-2001; 2001MO-JP01381.
 XX 24-FEB-2000; 2000JP-0047405.
 XX (KYOWA) KYOWA HAKKO KOGYO KK.
 PA (KITA) KITASATO INST.
 PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
 DR WPI; 2001-582053/65.
 XX N-PSDB; AAH92278.
 XX New modified avermectin aglycon synthase derived from Streptomyces
 PT avermectin used in production of 22,23-dihydroavermectin B1a used in
 PT drugs and pesticides -
 XX Claim 4; Page 217-235; 257pp; Japanese.

XX The present invention relates to the production of modified derivatives
 CC of avermectin aglycon synthase (AAS) derived from Streptomyces
 CC avermectin. The activity of an acyl carrier protein (ACP),
 CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
 CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is an S. avermectin protein.

XX Sequence 5532 AA;
 SQ Query Match 100.0%; Score 882; DB 22; Length 5532;
 Best Local Similarity 100.0%; Pred. No. 2.7e-76;
 Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSVLVTGCTGVLGAARHLAGVCGVRDLVSRGPDAPGAGLRAELAAAGAEVRIYA 60
 DB 1488 GSVLVTGCTGVLGAARHLAGVCGVRDLVSRGPDAPGAGLRAELAAAGAEVRIYA 1547

QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVFAAKVDALLLD 120

Db	1548	CDVGRRRVRLBEVVPAGCPLTGVVHAAGVLDATTISLPERGTFPAKVDAALLLD	1607
Oy	121	ELTRGMEISAFVLPSSAGTILGSAGCGNYAANAALDALAVRRRAAGLPGVSLAMGLW	178
Db	1608	ELTRGMEISAFVLPSSAGTILGSAGCGNYAANAALDALAVRRRAAGLPGVSLAMGLW	1665
RESULT 6			
AAAB23750			
ID	AAAB23750	standard; Protein; 6239 AA.	
XX	AAAB23750;		
XX			
DT	10-JAN-2001	(first entry)	
DE	S. avermectilis avermectin aglycon synthase protein SEQ ID NO:4.		
XX			
XX	Streptomyces avermectilis; avermectin aglycon synthase; biosynthesis;		
KW	multifunctional enzyme; polyketide; avermectin; veterinary drug;		
KM	agrochemical.		
XX	Streptomyces avermectilis.		
OS			
PN	WO200050605-A1.		
XX			
PD	31-AUG-2000.		
XX			
PF	23-FEB-2000; 2000WO-JP01041.		
XX			
PR	24-FEB-1999; 99JP-0046961.		
XX			
PA	(KITA) KITASATO INST.		
XX			
P1	Omura S, Ikeda H;		
DR	WPI; 2000-565458/52.		
DR	N-PSDB; AAA92301.		
PT	Avermectin aglycone synthase DNA and proteins encoded by all or part of		
PT	it for the production of avermectin and its derivatives for drug and		
PT	agrochemical use		
XX			
PS	Claim 32; Page 223-254; 314pp; Japanese.		
XX			
CC	The present sequence represents an avermectin aglycon synthase protein.		
CC	Also described are: (1) polypeptides encoded by all or part of the DNA;		
CC	(2) expression vectors containing the DNA; (3) host cells transformed by		
CC	the vectors; (4) preparation of the polypeptides by culture of the		
CC	transformants; (5) preparation of avermectin aglycon or its derivatives		
CC	by culture of transformed avermectin-producing microorganisms; and (6)		
CC	oligonucleotides of 5-60 bases in length containing sense or antisense		
CC	sequences from the avermectin aglycon synthase DNA. The enzymes are		
CC	useful for the production of modified forms of avermectin and of the		
CC	intermediates in its biosynthesis, for use as drugs, veterinary drugs		
CC	and agrochemicals.		
XX			
SQ	Sequence 6239 AA;		
Query Match	69.8%; Score 616; DB 21; Length 6239;		
Best Local Similarity	69.1%; Pred. No. 2.4e-50;		
Matches 123; Conservative 16; Mismatches 33; Indels 0; Gaps 0;			
Oy	1	GSVIVTGGTGVGAARVAAHLAGVCGVRDLLVSRGPDPAAGELRAELALGAERVIVA	60
Db	5753	GTIVLTGGTGVLRVLAHLEAHGVRRLLLAGRRGPABEAPLRAELGGATVEVVA	5812
Oy	61	CDVGRRRVRLBEVVPAGCPLTGVVHAAGVLDATTASLTPRLGTVFAAKVDALLLD	120
Db	5813	CDADRDQDLADLRLRPDDRPLTGVVHSAGILDGVTISLPERLGAVLRAKADAALLLD	5872
Oy	121	ELTRGMEISAFVLPSSAGTILGSAGCGNYAANAALDALAVRRRAAGLPGVSLAMGLW	178
Db	5873	ELTRGMEISAFVLPSSAGTILGSAGCGNYAANAALDALAVRRRAAGLPGVSLAMGLW	5930

ID	AA65265	standard; Protein; 6239 AA.
XX	AA65265;	
AC	04-DEC-2001	(first entry)
XX		
DT	Streptomyces avermiltillis protein SEQ ID NO: 5.	
XX		
DE	Avermectin aglycone synthase; AAS; avermectin derivative;	
KM	drug production; veterinary drug; pesticide.	
KW	Streptomyces avermiltillis.	
OS	WO200162939-A1.	
XX		
PN	30-AUG-2001.	
PD		
XX	23-FEB-2001; 2001WO-JP01381.	
PF		
PR	24-FEB-2000; 2000JP-0047405.	
XX		
PA	(KITOW) KITOWA HAKKO KOGYO KK.	
PA	(KITA) KITASATO INST.	
PI	Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;	
DR	N-PSDB; AAH79277.	
PT	New modified avermectin aglycone synthase derived from Streptomyces	
PT	avermiltillis used in production of 22,23-dihydroavermectin B1a used in	
PR	drugs and pesticides -	
PS	Claim 4; Page 180-201; 257dp; Japanese.	
CC	The present invention relates to the production of modified derivatives	
CC	of avermectin aglycone synthase (AAS) derived from Streptomyces	
CC	avermiltillis. The activity of an acyl carrier protein (ACP),	
CC	beta-ketoadacyl carrier protein synthase (KS), acyltransferase (AT),	
CC	beta-ketoadacyl carrier protein reductase (KR), dehydratase (DH), enoyl	
CC	reductase (ER) and/or thioesterase (TE) domain may be reduced or	
CC	suppressed. The process can be used in the production of drugs, veterinary	
CC	drugs and pesticides. The present sequence is an S. avermiltillis protein.	
SQ	Sequence 6239 AA;	
Query Match	69.8%; Score 616; DB 22; Length 6239;	
Best local Similarity	69.1%; Pred. No. 2,4e-50;	
Matches 123; Conservative	16; Mismatches 39; Indels 0; Gaps 0;	
OY	1 GSVLVITGSGTGVGAAGVARNHLAVGCVVDLLIVSRGPDAEGALRAELALGAENVIVA 60	
Db	5753 GTVLITGTGTGLVRLVARNHLYEANGVNHLILLAGRGSDAGAPELRALDLGGLTGEVEVA 5872	
OY	61 CDVGERRREVRRLLEGVPAGCPELTGVNNAAGVLDATTASTPERLGTVPFAKDADALLD 120	
Db	5813 CDADRDQQLDLRLRPDDRPLTGVSASGILLDDGVITTSIPSELTGLVLRKDKDALLD 5872	
OY	121 ELTGMELSAFVLFPSSAAGILGSAGCGNYAANAALDALVYRRRAAGLPGVSLAMGLM 178	
Db	5873 ELTRGALSFAVMSSASAVVGSFGCGNTYAANAANVLDPLAHRRRAEGLPVAVSLAMGLM 5930	
RESULT 8		
AA652718		
ID	AA652718	standard; Protein; 3724 AA.
MC	AA652718;	
XX		

ID	AA65265	standard; Protein; 6239 AA.
XX	AA65265;	
AC	04-DEC-2001	(first entry)
XX		
DT	Streptomyces avermiltillis protein SEQ ID NO: 5.	
XX		
DE	Avermectin aglycone synthase; AAS; avermectin derivative;	
KM	drug production; veterinary drug; pesticide.	
KW	Streptomyces avermiltillis.	
OS	WO200162939-A1.	
XX		
PN	30-AUG-2001.	
PD		
XX	23-FEB-2001; 2001WO-JP01381.	
PF		
PR	24-FEB-2000; 2000JP-0047405.	
XX		
PA	(KITOW) KITOWA HAKKO KOGYO KK.	
PA	(KITA) KITASATO INST.	
PI	Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;	
DR	N-PSDB; AAH79277.	
XX		
PT	New modified avermectin aglycone synthase derived from Streptomyces	
PT	avermiltillis used in production of 22,23-dihydroavermectin B1a used in	
PR	drugs and pesticides -	
PS	Claim 4; Page 180-201; 257bp; Japanese.	
CC	The present invention relates to the production of modified derivatives	
CC	of avermectin aglycone synthase (AAS) derived from Streptomyces	
CC	avermiltillis. The activity of an acyl carrier protein (ACP),	
CC	beta-ketoadacyl carrier protein synthase (KS), acyltransferase (AT),	
CC	beta-ketoadacyl carrier protein reductase (KR), dehydratase (DH), enoyl	
CC	reductase (ER) and/or thioesterase (TE) domain may be reduced or	
CC	suppressed. The process can be used in the production of drugs, veterinary	
CC	drugs and pesticides. The present sequence is an S. avermiltillis protein.	
SQ	Sequence 6239 AA;	
Query Match	69.8%; Score 616; DB 22; Length 6239;	
Best local Similarity	69.1%; Pred. No. 2,4e-50;	
Matches 123; Conservative	16; Mismatches 39; Indels 0; Gaps 0;	
OY	1 GSVLVITGCTGTGVAAGVARNHLAVGVNDLLIVSRGPDAEGALRAELALGAENVIVA 60	
Db	5753 GTVLITGCTGTGLVARHLVEAHGVNHLILLAGRGSDAGAPRLRALGLGATGEVVA 5812	
OY	61 CDVGERRRVRRLLEGVPAGCPLTGVNAAGAULDDATTASTPERLTGVFPAKDALLLD 120	
Db	5813 CDADRDQQLDLRLRPDPRPLTGVSASGILLDDGVITTSPELGLVLRKDKDALLLD 58722	
OY	121 ELTGMELSAFLVFSSAAGILGSAGGNVAAAANAALDALVYRRRAAGLPGVSLAMGLM 178	
Db	5873 ELTRGALSFAVMSSASAVVSGGNGNYAAAAANAVLDFLAHRRRAEGLPAVSLAMGLM 5930	
RESULT 8		
AA23718		
ID	AA23718 standard; Protein; 3724 AA.	
MC	AA23718;	
XX		

DT 27-FEB-1998 (first entry)
 XX Platenolide synthase ORF3 protein.
 DE
 XX Platenolide synthase gene cluster; platenolide production; ermG gene;
 KM multi-functional protein; macrocyclic antibiotic; spiramycin.
 XX
 OS Streptomyces ambifaciens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note="encoded by TGT"
 FT 36..459
 FT /note="ketosynthase domain, KS4"
 FT 594..921
 FT /note="acyltransferase domain, AT4"
 FT 1177..1350
 FT /note="ketosynthase domain, KS04"
 FT 1459..1523
 FT /note="acyl carrier protein domain, ACP4"
 FT 1545..1969
 FT /note="ketosynthase domain, KS5"
 FT 2088..2424
 FT /note="acyltransferase domain, AT5"
 FT 2445..2619
 FT /note="dehydratase domain, DHS"
 FT 2912..3261
 FT /note="enoylreductase domain, ERS"
 FT 3266..3451
 FT /note="ketoreductase domain, KRS"
 FT 3551..3631
 FT /note="acyl carrier protein domain, ACP5"
 FT Domain
 XX
 XX EP791656-A2.
 XX
 PD 27-AUG-1997.
 XX
 XX 19-FEB-1997; 97EP-0301066.
 XX
 XX 22-FEB-1996; 96US-0012050.
 XX
 PA (ELIL) LILLY & CO ELI.
 PI BURGESS SG, KUHSTOSS SA, RAO RN, RICHARDSON MA;
 PI ROSTECK PR;
 DR WPI, 1997-418047/39.
 DR N-PSDB; AAT78508.
 XX
 PT DNA encoding Streptomyces ambifaciens platenolide synthase domain -
 PT for production of spiramycin-related polyketide antibiotic
 XX
 PS Claim 8, Pages 54-66; 81p; English.
 XX
 CC AAM23716-W23720 represent proteins encoded by the platenolide synthase
 CC gene cluster of the invention. The gene cluster is also referred to as
 CC the ermG gene, and was isolated from Streptomyces ambifaciens. These
 CC sequences are multi-functional proteins which direct the synthesis of
 CC the polyketide platenolide. Platenolide is the basic building block of
 CC the macrocyclic antibiotic spiramycin. The DNA can be used to produce
 CC compounds exhibiting antibiotic activity based on the platenolide
 CC structure, including specifically the macrocyclic antibiotic spiramycin and
 CC spiramycin analogues and derivatives. Modifications of the platenolide
 CC synthase DNA sequence can be made so as to change the number and type of
 CC carboxylic acids incorporated into the growing polyketide chain and to
 CC change the kind of post-condensation processing that is conducted.
 XX
 SQ Sequence 3724 AA;

Query Match 64.4%; Score 568; DB 18; Length 3724;
 Best Local Similarity 65.2%; Pred. No. 6.2e-46;
 Matches 116; Conservative 19; Mismatches 43; Indels 0; Gaps 0;

QY 1 GSVLVTGTTGTVLGAANVAHLAGVGVRLDLLVSRGRPDAPGAGLRAELALGAEVRIYA 60
 DB 3268 GTVLVTGTGTGALGSRITARHLASRGVRLILAAARRGPDEGAALVADIALGASATVVA 3327
 QY 61 CDVGERREVRRLLEGPACCEPLTGVVHAAGVDDATIAITPERLGTVPAAKVDALLLD 120
 DB 3328 CDVSDADAVRGLAGIPADHPITAVVHSYGVDDGVLPGLTPEKMRVLRPKVEAAVHLD 3387
 QY 121 ELTRGMELSAFVLFSSAGILGSAGCGNVAANAAALDALAYRRRAGLPGVSLANGIM 178
 DB 3388 ELTRDUDLSAFVLFSSAGILGSAGCGNVAANAAATIDALAAARRSLGLPSVSLANGIM 3445
 RESULT 9
 ID AAM22608 standard; protein; 3724 AA.
 XX
 AC AAM22608;
 XX
 DT 27-FEB-1998 (first entry)
 XX
 DE Platenolide synthase ORF3 protein.
 XX
 KM Tyactone synthase gene cluster; tylG gene; multifunctional protein;
 KM polyketide; tyactone synthase; antibiotic; tylosin.
 XX
 OS Streptomyces ambifaciens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note="encoded by TGT"
 FT 36..459
 FT /note="ketosynthase domain, KS4"
 FT 594..921
 FT /note="acyltransferase domain, AT4"
 FT 1177..1350
 FT /note="ketosynthase domain, KS04"
 FT 1459..1523
 FT /note="acyl carrier protein domain, ACP4"
 FT 1545..1969
 FT /note="ketosynthase domain, KS5"
 FT 2088..2424
 FT /note="acyltransferase domain, AT5"
 FT 2445..2619
 FT /note="dehydratase domain, DHS"
 FT 2912..3261
 FT /note="enoylreductase domain, ERS"
 FT 3266..3451
 FT /note="ketoreductase domain, KRS"
 FT 3551..3631
 FT /note="acyl carrier protein domain, ACP5"
 FT Domain
 XX
 XX EP791655-A2.
 XX
 PD 27-AUG-1997.
 XX
 XX 19-FEB-1997; 97EP-0301056.
 XX
 XX 22-FEB-1996; 96US-0012078.
 XX
 PA (ELIL) LILLY & CO ELI.
 PI Dehoff BS, Kuhstoss SA, Rostock PR, Sutton KU;
 DR WPI, 1997-418046/39.
 DR N-PSDB; AAT80414.
 XX
 PT DNA encoding Streptomyces fradiae tyactone synthase domain - for
 PT production of tylosin-related polyketide compounds
 XX
 PS Example 2; Pages 155-167; 220pp; English.
 XX
 CC AAM22606-W22610 represent proteins encoded by the platenolide synthase

PT	Domain	35..455	/label= KS3_domain
PT		/note="Ketosynthase (KS) domain"	
PT	Domain	546..858	/label= AT3_domain
PT		/note="Acyltransferase (AT) domain"	
PT	Domain	892..1073	/label= DH3_domain
PT		/note="Dehydratase (DH) domain"	
PT	Domain	1381..1628	/label= KR3_domain
PT		/note="Ketoreductase (KR) domain"	
PT	Domain	1662..1735	/label= ACP3_domain
PT		/note="Acyl carrier protein (ACP) domain"	
PT	Domain	1757..2180	/label= KS4_domain
PT		/note="Ketosynthase (KS) domain"	
PT	Domain	2291..2603	/label= AT4_domain
PT		/note="Acyltransferase (AT) domain"	
PT	Domain	2617..2818	/label= DH4_domain
PT		/note="Dehydratase (DH) domain"	
PT	Domain	3124..3371	/label= KR4_domain
PT		/note="Ketoreductase (KR) domain"	
PT	Domain	3407..3480	/label= ACP4_domain
PT		/note="Acyl carrier protein (ACP) domain"	
PT	Domain	3501..3924	/label= KS5_domain
PT		/note="Ketosynthase (KS) domain"	
PT	Domain	4032..4346	/label= AT5_domain
PT		/note="Acyltransferase (AT) domain"	
PT	Domain	4360..4561	/label= DH5_domain
PT		/note="Dehydratase (DH) domain"	
PT	Domain	4953..5239	/label= ER5_domain
PT		/note="Enoylreductase (ER) domain"	
PT	Domain	5248..5495	/label= KR5_domain
PT		/note="Ketoreductase (KR) domain"	
PT	Domain	5528..5601	/label= ACP5_domain
PT		/note="Acyl carrier protein (ACP) domain"	
PT	Domain	5623..6046	/label= KS6_domain
PT		/note="Ketosynthase (KS) domain"	
PT	Domain	6166..6478	/label= AT6_domain
PT		/note="Acyltransferase (AT) domain"	
PT	Domain	6492..6704	/label= DH6_domain
PT		/note="Dehydratase (DH) domain"	
PT	Domain	7038..7281	/label= KR6_domain
PT		/note="Ketoreductase (KR) domain"	
PT	Domain	7315..7388	/label= ACP6_domain
PT		/note="Acyl carrier protein (ACP) domain"	
PT	Domain	7408..7831	/label= KS7_domain
PT		/note="Ketosynthase (KS) domain"	
PT	Domain	7939..8253	/label= AT7_domain
PT		/note="Acyltransferase (AT) domain"	
PT	Domain	8267..8470	/label= DH7_domain
PT		/note="Dehydratase (DH) domain"	
PT	Domain	8812..9086	

Query Match	Best Local Similarity	Score	DB	Length	11096
Matches 116; Conservative 12; Mismatches 50; Indels 0; Gaps 0;					
1 GSVLVTGGTGTGLGAARHLAGVGVGRDILLVSRGPDAPGAEGLRAELALGAEVRIVA 60					
8812 GTVLTITGGTGLGQPARHLVDRGVGRNLLVSRGPDAPETTELVAELTHGAEVAVQA 8871					
61 CDVGERREVRALLGCVAGCPLTGVVAAGVLDATIASLTPERLQTVFAKVDALLLD 120					
8872 CDVADGVAALVAGVDEPHPLRAVVTAGVLDGVGSLTBEELATVLRKADAAWHIH 8931					

QY 121 ELTRGELSAFVLFSSAAGILGSAGCGNTAANAALDALAYRRRAAGLPGVSLAMGLW 178
 DB 8932 EATGGLDPAFVTVSSVAGVFGAGQANYAANAFLDLMAGRRALPGSLAMGLW 8989

RESULT 12

ID AAB10144 standard; Protein, 9477 AA.

XX AAB10144;

XX 29-NOV-2001 (first entry)

DE Streptomycetes noursei nystatin gene, NysI (complete).

KM Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;

KW antifungal; antibiotic; PKS type I.

XX Streptomycetes noursei.

XX Location/Qualifiers

FT 41..464

FT /label= KS15 domain

FT /note= "Ketosynthase (KS) domain"

FT 578..889

FT /label= AT15 domain

FT /note= "Acyltransferase (AT) domain"

FT 903..1102

FT /label= DH15 domain

FT /note= "Dehydratase (DH) domain"

FT 1446..1731

FT /label= KR15 domain

FT /note= "Enoylreductase (ER) domain"

FT 1740..1988

FT /label= KR15 domain

FT /note= "Ketoerectase (KR) domain"

FT 2023..2096

FT /label= ACP15 domain

FT /note= "Acyl carrier protein (ACP) domain"

FT 2117..2538

FT /label= KS16 domain

FT /note= "Ketosynthase (KS) domain"

FT 2635..2953

FT /label= AT16 domain

FT /note= "Acyltransferase (AT) domain"

FT 2967..3167

FT /label= DH16 domain

FT /note= "Dehydratase (DH) domain"

FT 3257..3500

FT /label= KR16 domain

FT /note= "Ketoerectase (KR) domain"

FT 3539..3612

FT /label= ACP16 domain

FT /note= "Acyl carrier protein (ACP) domain"

FT 3634..4057

FT /label= KS17 domain

FT /note= "Ketosynthase (KS) domain"

FT 4153..4472

FT /label= AT17 domain

FT /note= "Acyltransferase (AT) domain"

FT 4486..4725

FT /label= DH17 domain

FT /note= "Dehydratase (DH) domain"

FT 4997..5245

FT /label= KR17 domain

FT /note= "Ketoerectase (KR) domain"

FT 5277..5350

FT /label= ACP17 domain

FT /note= "Acyl carrier protein (ACP) domain"

PD 16-AUG-2001.

XX 08-FEB-2001; 2001MO-GB00509.

XX 08-FEB-2000; 2000GB-0002840.

PR 10-APR-2000; 2000GB-0008786.

PR 14-APR-2000; 2000GB-0009387.

XX (UYNO-) UNIV NORGES TEKNIKSKAPETELIGE.

PA (SNTF) SINTEF STIFTELSEN IND TEK FORSK.

PA (ALPH-) ALPHARMA AS.

PA (SINV-) SINVENT AS.

PA (DZIE/) DZIELEWSKA H.

PA (ZOTC/) ZOTCHEV S B.

PA (SEKU/) SEKUROVA O N.

PA (FJAE/) FJAEVIT E.

PA (BRAU/) BRAUTASET T.

PA (STRO/) STROM A R.

XX Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;

PI Valla S, Ellingsen TE, Sletta H, Gulliksen O;

XX WPI, 2001-557614/62.

DR N-PSDB; AAD17186.

XX New nystatin polyketide synthase polynucleotides and polypeptides,

PT useful as antibiotics and antifungals.

XX Claim 15; Page 255-260; 266pp; English.

PS The present invention relates to the cloning and sequencing of the gene

XX cluster encoding a modular type I polyketide synthase (PKS) enzyme

CC involved in the biosynthesis of the macrolide antibiotic nystatin.

CC The nystatin PKS is useful as antifungal antibiotic. The present

CC sequence is a PKS type I encoding Streptomycetes noursei

CC nystatin gene, NysI (complete).

XX Sequence 9477 AA;

SO Query Match 62.2%; Score 548.5; DB 22; Length 9477;

XX Best Local Similarity 63.5%; Pred. No. 1.5e-43;

XX Matches 113; Conservative 21; Mismatches 43; Indels 1; Gaps 1;

QY 1 GSVLTGTVGTVGAARHLAGVGVDDLLVRRGPDPAAGALRAELAAAGVRIYA 60

DB 9050 GIVLTGTVGTVGAARHLAGVGVDDLLVRRGPDPAAGALRAELAAAGVRIYA 9108

QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPELGTVPAAKYDALILD 120

DB 9109 CDVADRDQTLAVLDALPADRPPLTGVVHTAGVLDGVLDRTPERFQEVFRAKVTSLILD 9168

QY 121 ELTRGELSAFVLFSSAAGILGSAGCGNTAANAALDALAYRRRAAGLPGVSLAMGLW 178

DB 9169 ELTRRELAFAVLFSSAAGVNPQANYAANAALDALAQRVYGLPATSVSGAW 9226

RESULT 13

ID AAB18638 standard; Protein, 3739 AA.

XX AAB18638;

XX 22-JUN-2001 (first entry)

XX Amino acid sequence of narbonolide synthase subunit 2 (P1CA11).

DE Narbonolide synthase; polyketide synthase gene; narbonolide polyketide;

XX antibiotic; C12-hydroxylase; pick; desosamine biosynthesis;

KM desosaminyl transferase enzyme; ketolide; beta-glucosidase enzyme;

KW picrotoxin biosynthesis.

XX Streptomycetes venezuelae.

XX Streptomycetes venezuelae.

XX Streptomycetes venezuelae.

PN US6117659-A.
XX 12-SEP-2000.
XX
XX
XX 27-MAY-1999; 99US-0320878.
XX
XX 28-MAY-1998; 98US-0087080.
XX 22-SEP-1998; 98US-0100880.
XX 08-FEB-1999; 99US-0119139.
XX 20-MAY-1999; 99US-0134980.
XX 30-APR-1997; 97US-0846247.
XX 06-MAY-1998; 98US-0073538.
XX 28-AUG-1998; 98US-0141908.
XX
XX (KOSA-) KOSAN BIOSCIENCES INC.
XX
XX Ashley G, Beclach MC, Beclach M, Tang L, McDaniel R;
XX WPI, 2000-610844/58.
XX
XX New recombinant plck hydroxylase gene of Streptomyces venezuelae useful
XX for converting ketolides to antibiotics and
XX intermediates in the synthesis of compounds with pharmaceutical value
XX
XX Disclosure; Columns 11-12; 117pp; English.
XX
XX The present sequence represents a narbonolide synthase subunit 2
XX (PICK11). The nucleotide sequence encoding it is used in the course of
XX the invention. The specification describes a recombinant DNA compound
XX expressing recombinant polypeptide synthase genes in host cells for the
XX production of narbonolide, narbonolide derivatives and polypeptides that
XX are useful as antibiotics and as intermediates in the synthesis of
XX compounds with pharmaceutical value. The DNA compounds may also encode
XX a C12-hydroxylase (pick), desosamine biosynthesis and desosaminyl
XX transferase enzymes (useful for conversion of ketolides to antibiotics),
XX and the beta-glucosidase enzyme (involved in picromycin biosynthesis).
XX These compounds are also useful for increasing the antibiotic activity
XX of a compound relative to the unhydroxylated compound. The recombinant
XX host cells are useful as genetic systems that allow rapid engineering
XX of the narbonolide polypeptide synthase. These would be valuable for
XX creating novel ketolide analogs for pharmaceutical applications.
XX
XX Sequence 3739 AA;
SQ
Query Match 62.0%; Score 547; DB 21; Length 3739;
Best Local Similarity 62.8%; Pred. No. 6.9e-44;
Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;
QY 1 GSVLVTGGTGVGAANVARRHLAGVCGVNDLLVSRGGPDAGAGELRAELAAAGAEVRIYA 60
DB 3277 GTVLLTGGTGAAGIYARRHVGKVRRLVSRGGPDAGAGELVHELALGADSVAA 3336
QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDAALLD 120
DB 3337 CDVADREALTAVLDSPHAPHPLTAVVHTAGVLSGTLPSMTAEDVEHVLRPKVDAAFILD 3396
QY 121 ELTR--GMELSAVLFPSSAAGILGSAGCGNYAANAALDLARRRAAGIPGVSLAWGLW 178
DB 3397 ELTSTPGYDLAAVFMSSAAAVFGAGGAGVAAANATLDALAWRRRTAGLPALSLGGLW 3456

KW neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
KW biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;
KW chronic obstructive pulmonary disease; respiratory inflammation;
KW hypercholesterolaemia; crop protection agent.
OS Streptomyces venezuelae ATCC15439.
XX
XX WO200000620-A2.
XX
XX 06-JUN-2000.
XX
XX 25-JUN-1999; 99WO-US14398.
XX
XX 26-JUN-1998; 98US-0105537.
XX
XX (MIND) UNIV MINNESOTA.
XX
XX Sherman DH, Liu H, Xue Y, Zhao L;
XX WPI, 2000-160679/14.
XX
XX N-PSDB; AA287298.
XX
XX Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.
XX synthesis of methymycin and pikromycin -
XX
XX Claim 19; Page 403-415; 438pp; English.
XX
XX The invention relates to an isolated and purified nucleic acid segment
XX comprising a desosamine biosynthetic gene cluster, a fragment or its
XX biologically active variant, where the nucleic acid sequence is not
XX derived from the eryc gene cluster of Saccharopolyspora erythraea or
XX Streptomyces antibioticus. The invention also relates to a macrolide
XX biosynthetic gene cluster, or fragments thereof, the macrolide
XX biosynthetic gene cluster encodes proteins which synthesize methymycin,
XX pikromycin, neomethymycin, narbomycin or a combination of these
XX compounds. Recombinant or augmented cells comprising the desosamine
XX and/or macrolide biosynthetic gene clusters are useful for the
XX production of biologically active macrolides. The macrolide biosynthetic
XX production are useful for synthesis of methymycin, pikromycin,
XX neomethymycin and narbomycin. The alternative termination of polypeptide
XX synthesis may be useful to prepare novel antibiotics and
XX polyhydroxyalkanoate (PHA) monomers. The compounds produced by the
XX recombinant host cells are useful as biopolymers, e.g., in packaging or
XX biomedical applications, to engineer PHA monomer synthases or to prepare
XX biologically active agents, such as chemotherapeutics.
XX immunosuppressants, agents to treat asthma, chronic obstructive pulmonary
XX disease as well as other diseases involving respiratory inflammation,
XX cholesterol-lowering agents or macrolide-based antibiotics which are
XX active against a variety of organisms, e.g., bacteria, including
XX multi-drug resistant pneumococci and other respiratory pathogens, as well
XX as viral parasitic pathogens, or as crop protection agents (e.g.,
XX fungicides or insecticides) via expression of polypeptides in plants.
XX Sequences AA77190-Y77197 represent macrolide biosynthetic enzymes from
XX Streptomyces venezuelae ATCC 15439, which are encoded by sequences
XX AA287295-287302.
XX
XX Sequence 3739 AA;
SQ
Query Match 62.0%; Score 547; DB 21; Length 3739;
Best Local Similarity 62.8%; Pred. No. 6.9e-44;
Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;
QY 1 GSVLVTGGTGVGAANVARRHLAGVCGVNDLLVSRGGPDAGAGELRAELAAAGAEVRIYA 60
DB 3277 GTVLLTGGTGAAGIYARRHVGKVRRLVSRGGPDAGAGELVHELALGADSVAA 3336
QY 61 CDVGERREVRRLLEGVPACPLTGVVHAAGVLDATIASLTPERLGTVPAAKVDAALLD 120
DB 3337 CDVADREALTAVLDSPHAPHPLTAVVHTAGVLSGTLPSMTAEDVEHVLRPKVDAAFILD 3396
QY 121 ELTR--GMELSAVLFPSSAAGILGSAGCGNYAANAALDLARRRAAGIPGVSLAWGLW 178
DB 3397 ELTSTPGYDLAAVFMSSAAAVFGAGGAGVAAANATLDALAWRRRTAGLPALSLGGLW 3456

RESULT 15

AAV77201
ID AAV77201 standard; Protein; 3739 AA.

AAV77201;

05-JUN-2000 (first entry)

S. venezuelae pik (macrolide biosynthesis) gene cluster protein #2.

Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;
chronic obstructive pulmonary disease; respiratory inflammation;
hypercholesterolemia; crop protection agent.

Streptomyces venezuelae ATCC15439.

WC200000620-A2.

06-JAN-2000.

25-JUN-1999; 99WO-US14398.

26-JUN-1998; 98US-0105537.

(MINU) UNIV MINNESOTA.

Sherman DH, Liu H, Xue Y, Zhao L;

WPI; 2000-160679/14.

DR N-PSDB; AA287318.

Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.

synthesis of methymycin and pikromycin -

PS Disclosure; Figure 31; 438bp; English.

The invention relates to an isolated and purified nucleic acid segment comprising a desosamine biosynthetic gene cluster, a fragment or its biologically active variant, where the nucleic acid sequence is not derived from the eryC gene cluster of Saccharopolyspora erythraea or Streptomyces antibioticus. The invention also relates to a macrolide biosynthetic gene cluster, or fragments thereof. The macrolide biosynthetic gene cluster encodes proteins which synthesise methymycin, pikromycin, neomethymycin, narbomycin or a combination of these compounds. Recombinant or augmented cells comprising the desosamine and/or macrolide biosynthetic gene clusters are useful for the production of biologically active macrolides. The macrolide biosynthetic proteins are useful for synthesis of methymycin, pikromycin, neomethymycin and narbomycin. The alternative termination of polyketide synthesis may be useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA) monomers. The compounds produced by the recombinant host cells are useful as biopolymers, e.g., in packaging or biomedical applications, to engineer PHA monomer synthases or to prepare biologically active agents, such as chemotherapeutics, immunosuppressants, agents to treat asthma, chronic obstructive pulmonary disease as well as other diseases involving respiratory inflammation, cholesterol-lowering agents or macrolide-based antibiotics which are active against a variety of organisms, e.g., bacteria, including multi-drug resistant pneumococci and other respiratory pathogens, as well as viral parasitic pathogens, or as crop protection agents (e.g., fungicides or insecticides) via expression of polyketides in plants. Sequences AAV77200-Y77203 and AAV80997 represent proteins encoded by the macrolide biosynthetic gene cluster (pik) from Streptomyces venezuelae ATCC 15439, as given in figure 31.

Sequence 3739 AA;

Query Match 62.0%; Score 547; DB 21; Length 3739;
Best Local Similarity 62.8%; Pred. No. 6,9e-44;

Matches 113; Conservative 19; Mismatches 46; Indels 2; Gaps 1;

Qy	1	GSVLVNGGTVGA	AAV77201	AGVGRDILLVSRGGPDAPGAGLAAELAAEAERYIVA	60
Db	3277	GIYLLTGGTGAL	GIYARHVGEMVGRLLVSRGGTDAPGAGELVHELEALGADVSVA	3336	
Qy	61	CDVGERREVRRL	EGVPAGCPLTGVAAGVLDATIASLTPERLGTVPAAKYDAALLD	120	
Db	3337	CDVADREALTAV	LDSPAEHPLTAVVHTAGVSDGTLPSVTADVEHVRPKVDAAFLD	3396	
Qy	121	ELTR--GMEISAV	LVFSSAAGILGSGCGENYAAANALDALAYRRPAAGLPGVSLAWGLW	178	
Db	3397	ELTSTPGYDIAA	RVFMFSSAAAVGAGGQGYAANAATLDAWRRRTAGLPALSLGWLW	3456	

Search completed: June 17, 2003, 13:02:34
Job time : 9.71975 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:53:02 ; Search time 14.6048 Seconds
(without alignments)
4458.180 Million cell updates/sec

Title: US-09-914-286-3_COPY_29_344

Perfect score: 1618
Sequence: 1 VFVPCGPGQPMGREGRLD.....ADTVINGTLRGGTLDHFL 316

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1618	100.0	3972	2	Q9S0R8	Q9S0R8 streptomyc
2	1578	97.5	3626	2	Q9F779	Q9F779 streptomyc
3	819	50.6	3654	2	Q30766	Q30766 streptomyc
4	817.5	50.5	8817	2	Q53840	Q53840 polyanthum
5	791	48.9	4881	2	Q9S0R3	Q9S0R3 streptomyc
6	773	47.8	4151	16	O51490	O51490 mycobacteri
7	765.5	47.3	3192	2	Q9L4M4	Q9L4M4 streptomyc
8	763	47.3	2595	2	Q9ALM6	Q9ALM6 saccharopol
9	763	47.2	3939	2	Q93HJ3	Q93HJ3 streptomyc
10	762	47.1	3729	2	Q33956	Q33956 streptomyc
11	758	46.8	3970	2	Q93HJ8	Q93HJ8 streptomyc
12	756.5	46.8	4498	2	Q93HE5	Q93HE5 streptomyc
13	756	46.7	9510	2	Q93HX9	Q93HX9 streptomyc
14	754.5	46.6	9507	2	Q9EMH1	Q9EMH1 streptomyc
15	753	46.5	9477	2	Q9L4X3	Q9L4X3 streptomyc
16	750.5	46.4	3190	2	Q93NM7	Q93NM7 streptomyc

17	750	46.4	5588	2	Q9ALM2	Q9ALM2 saccharopol
18	746	46.1	3562	2	Q9F829	Q9F829 microsomosp
19	745.5	46.1	1041	2	Q93C90	Q93C90 actinomdur
20	740	45.7	3352	2	Q93H83	Q93H83 streptomyc
21	737	45.6	4685	2	Q93HJ2	Q93HJ2 streptomyc
22	734	45.4	5532	2	Q9S0R4	Q9S0R4 streptomyc
23	733	45.3	3613	2	Q93HJ1	Q93HJ1 streptomyc
24	731	45.2	6146	2	Q93HJ5	Q93HJ5 streptomyc
25	722.5	44.7	6315	2	Q9ADL6	Q9ADL6 polyanthum
26	721	44.6	1907	2	Q93HJ9	Q93HJ9 streptomyc
27	718.5	44.4	3201	2	Q9F828	Q9F828 microsomosp
28	714.5	44.2	3816	2	Q9KIV3	Q9KIV3 streptomyc
29	713	44.1	6239	2	Q9S0R7	Q9S0R7 streptomyc
30	712	44.0	4150	2	Q9KIV4	Q9KIV4 streptomyc
31	706	43.6	4840	2	Q93HJ4	Q93HJ4 streptomyc
32	704	43.5	3546	2	Q9F830	Q9F830 microsomosp
33	699	43.2	1937	2	Q30482	Q30482 streptomyc
34	697.5	43.1	5060	2	Q52545	Q52545 amycolatops
35	697.5	43.1	5069	2	Q52789	Q52789 amycolatops
36	696.5	43.0	4613	2	Q92G15	Q92G15 streptomyc
37	689.5	42.6	2126	16	P94996	P94996 mycobacteri
38	685	42.3	3524	2	Q93H86	Q93H86 streptomyc
39	681.5	42.1	3739	2	Q92G14	Q92G14 streptomyc
40	674.5	41.7	3170	2	Q9ALM4	Q9ALM4 saccharopol
41	674	41.7	1562	2	Q92G13	Q92G13 streptomyc
42	673.5	41.6	1762	2	Q52546	Q52546 amycolatops
43	673.5	41.6	1763	2	Q52790	Q52790 amycolatops
44	672	41.5	7525	2	Q9KIE0	Q9KIE0 streptomyc
45	667.5	41.3	1346	2	Q92G12	Q92G12 streptomyc

ALIGNMENTS

RESULT 1

Q9S0R8 PRELIMINARY; PRT; 3972 AA.
ID Q9S0R8
AC Q9S0R8;
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 1.
GN AVEA1.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
OC NCBI_taxid=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9380548; PubMed=10449723;
RA Ireda H., Nonomiyama T., Uemura M., Ohta T., Omura S.,
RT "Organization of the biosynthetic gene cluster for the polyketide
anhydride macrolide avermectin in Streptomyces avermitilis.";
RT Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
RL EMBL; AB032367; BAA84474.1; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantoate attach.
DR InterPro; IPR000834; Zn cardopect.
DR Pfam; PF00698; Acyl_transf. 3.
DR Pfam; PF0109; ketoacyl-synt. 2.
DR Pfam; PF02801; ketoacyl-synt. C; 2.
DR Pfam; PF0550; pp-binding; 3.
DR PROSITE; PS50075; ACP DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL SYNTHASE; 2.
DR PROSITE; PS00133; CARBOXYPEPT. ZN 2; UNKNOWN_2.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 2.
KM Phosphopantetheine; Transferase.
SQ SEQUENCE 3972 AA; 416852 MW; 2A293695S032B1C3 CRC64;

Query Match 100.0%; Score 1618; DB 2; Length 3972;
Best Local Similarity 100.0%; Pred. No. 3.5e-100;
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 VFVPPGGPQWPGMGRRLDASDVFRSVRACAAAFAPYVDMVEOVLRSDDAPAGLDV 60
DB 29 VFVPPGGPQWPGMGRRLDASDVFRSVRACAAAFAPYVDMVEOVLRSDDAPAGLDV 88
QY 61 DVVQPTLFAVWISLAALMRSGVPCAVLGHSIGELIAAAHVSGSLADAAVVTLSQA 120
DB 89 DVVQPTLFAVWISLAALMRSGVPCAVLGHSIGELIAAAHVSGSLADAAVVTLSQA 148
QY 121 QTTLAGTALVSAATPDELLPRIAPWTEDEPARLAVAANVGRSTVSGAREAVADLVA 180
DB 149 QTTLAGTALVSAATPDELLPRIAPWTEDEPARLAVAANVGRSTVSGAREAVADLVA 208
QY 181 DLTAAGVRTMIIPVDVPAHSPLMTAIEBRVVSGLPTTPRPSRIIPHSVYTGRLDTREL 240
DB 209 DLTAAGVRTMIIPVDVPAHSPLMTAIEBRVVSGLPTTPRPSRIIPHSVYTGRLDTREL 268
QY 241 DAAVYRNMSSTVRFEPARILLLOQPKTFVEMSPHPVLTMGLOELAPDLGDTTGTADTV 300
DB 269 DAAVYRNMSSTVRFEPARILLLOQPKTFVEMSPHPVLTMGLOELAPDLGDTTGTADTV 328
QY 301 IMGTLRGGGCTLDHFL 316
DB 329 IMGTLRGGGCTLDHFL 344

```

RESULT 2

09F779 PRELIMINARY; PRT; 3626 AA.

```

AC 09F779;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Avermectin polyketide synthase (fragment).
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC31271;
RA Hong Y.-S., Lee J.-J.;
RT "Targeted Gene Disruption of the avermectin O-methyltransferase gene
RT and polyketide synthase gene from Streptomyces avermitilis.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF275943; AAC09812.1; -
DR InterPro: IPR001227; Ac_transferase.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR003880; Pantane_attach.
DR Pfam: PF00698; Acyl_transf. 3.
DR Pfam: PF00109; Ketoacyl-synt. 2.
DR Pfam: PF02801; ketoacyl-synt. C; 2.
DR Pfam: PF00550; pp-binding; 2.
DR PROSITE: PS50075; ACP_DOMAIN; 2.
DR PROSITE: PS00016; B_KETOACYL_SYNTHASE; 2.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
FT NON TER 3626
SQ SEQUENCE 3626 AA; 380557 MW; 6272F5F088C1A8D0 CRC64;

```

Query Match 97.5%; Score 1578; DB 2; Length 3626;
 Best Local Similarity 98.1%; Pred. No. 1.6e-97;
 Matches 310; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 VFVPPGGPQWPGMGRRLDASDVFRSVRACAAAFAPYVDMVEOVLRSDDAPAGLDV 60
DB 10 VFVPPGGPQWPGMGRRLDASDVFRSVRACAAAFAPYVDMVEOVLRSDDAPAGLDV 69
QY 61 DVVQPTLFAVWISLAALMRSGVPCAVLGHSIGELIAAAHVSGSLADAAVVTLSQA 120
DB 70 DVVQPTLFAVWISLAALMRSGVPCAVLGHSIGELIAAAHVSGSLADAAVVTLSQA 129
QY 121 QTTLAGTALVSAATPDELLPRIAPWTEDEPARLAVAANVGRSTVSGAREAVADLVA 180

```

```

DB 130 QTTLAGTALVSAATPDELLPRIAPWTEDEPARLAVAANVGRSTVSGAREAVADLVA 189
QY 181 DLTAAGVRTMIIPVDVPAHSPLMTAIEBRVVSGLPTTPRPSRIIPHSVYTGRLDTREL 240
DB 190 DLTAAGVRTMIIPVDVPAHSPLMTAIEBRVVSGLPTTPRPSRIIPHSVYTGRLDTREL 249
QY 241 DAAVYRNMSSTVRFEPARILLLOQPKTFVEMSPHPVLTMGLOELAPDLGDTTGTADTV 300
DB 250 DAAVYRNMSSTVRFEPARILLLOQPKTFVEMSPHPVLTMGLOELAPDLGDTTGTADTV 309
QY 301 IMGTLRGGGCTLDHFL 316
DB 310 IMGTLRGGGCTLDHFL 325

```

RESULT 3

030766 PRELIMINARY; PRT; 3654 AA.

```

AC 030766;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Polyketide synthase modules 4 and 5.
GN NID3.
OS Streptomyces caelestis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=36816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL-2821;
RX MEDLINE=98053867; Pubmed=9393718;
RA Kakavas S.J., Katz L., Stassi D.;
RT "Identification and characterization of the niddamycin polyketide
RT synthase genes from Streptomyces caelestis.";
RL J. Bacteriol. 179:7515-7522(1997).
DR EMBL: AF016585; AAC46026.1; -
DR InterPro: IPR001227; Ac_transferase.
DR InterPro: IPR002085; Adh_zn_family.
DR InterPro: IPR001827; Antennapedia.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR003880; Pantane_attach.
DR InterPro: IPR002364; OOR_zeta_crystal.
DR Pfam: PF00698; Acyl_transf. 2.
DR Pfam: PF00107; adh_zinc; 1.
DR Pfam: PF00109; ketoacyl-synt. 2.
DR Pfam: PF02801; ketoacyl-synt. C; 2.
DR Pfam: PF00550; pp-binding; 2.
DR PROSITE: PS50075; ACP_DOMAIN; 2.
DR PROSITE: PS00032; ANTENNAPEDIA; 1.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_2.
DR PROSITE: PS01162; OOR_ZETA_CRYSTAL; 1.
DR PHOSPHOPANTHETHEINE; Transferase.
FT NON TER 3654
SQ SEQUENCE 3654 AA; 379108 MW; 59825B49EDCD06FF CRC64;

```

Query Match 50.6%; Score 819; DB 2; Length 3654;
 Best Local Similarity 54.0%; Pred. No. 2.4e-46;
 Matches 168; Conservative 38; Mismatches 95; Indels 10; Gaps 3;

```

QY 1 VFVPPGGPQWPGMGRRLDASDVFRSVRACAAAFAPYVDMVEOVLRSDDAPAGLDV 60
DB 2008 VFVPPGGPQWPGMGRRLDASDVFRSVRACAAAFAPYVDMVEOVLRSDDAPAGLDV 2067
QY 61 DVVQPTLFAVWISLAALMRSGVPCAVLGHSIGELIAAAHVSGSLADAAVVTLSQA 120
DB 2068 DVVQPTLFAVWISLAALMRSGVPCAVLGHSIGELIAAAHVSGSLADAAVVTLSQA 2127
QY 121 QTTLAGTALVSAATPDELLPRIAPWTEDEPARLAVAANVGRSTVSGAREAVADLVA 180
DB 2128 WIGLAGKGWVAAPVMAEELRPLVTMGD-----RLAAVAVNSPGSCAVAGDPEALAEIYA 2183

```



```

QY 181 DLTAQVTRMIP-VDVPAHSPLMYAIEERVSGLLPTPRSRIPPHSSVTGRLDPTRE 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2184 LTGEGHARPIPEVDNAGHSPOVDALRAHLBYLAVAPRPADIPFYSTVGGLDGTE 2243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 240 LDAAYWYRNNSSTVFEPAPARLLIQGPKTFVEMSPHPVLTMGLOELAPDLGDTTGTADT 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2244 LDATYWRNMRPEVEFRATRALIADGHDFLETSPPHMLAVALAQVTADG-----TDA 2298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 300 VIMGLTARGOG 310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2299 AVLGTLRRRG 2309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
053840
ID 053840 PRELIMINARY; PRT; 8817 AA.
AC Q53840;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Soraphen polyketide synthase B.
GN SORB.
OS Polyangium celluloseum.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Sorangineae; Polyangiaceae; Polyangium.
OX NCBI_TaxID=56;
RN [1]
RP SEQUENCE OF 6342-8466 FROM N.A.
RC STRAIN=SO CE26;
RA MEDLINE=9535306; PubMed=7601830;
RA Schupp T., Toupet C., Cluzel B., Neff S., Hill S., Beck J.J.,
RA Lygon J.M.;
RA "A Sorangium celluloseum (myxobacterium) gene cluster for the
RT biosynthesis of the macrolide antibiotic soraphen A: cloning,
RT characterization, and homology to polyketide synthase genes from
RT actinomycetes."
RL J. Bacteriol. 177:3673-3679(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SO CE26;
RA Lygon J.M., Hill S., Beck J., Zirkle R., Molnar I., Zawodny J.,
RA Money S., Schupp T.;
RT "Characterization of the biosynthetic gene cluster for the antifungal
RT polyketide soraphen A from Sorangium celluloseum So CE26."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DB EMBL; U24241; AAA79984.2; -.
DR HSSP; P24249; 1HNK.
DR Interpro; IPR001227; Ac transferase.
DR Interpro; IPR002198; ADH_short.
DR Interpro; IPR002085; Adh_family.
DR Interpro; IPR004410; FAD.
DR Interpro; IPR000794; ketoacyl-synt.
DR Interpro; IPR003880; Pantne_attach.
DR Interpro; IPR003662; sub transporter.
DR Interpro; IPR001031; Thioesterase.
DR Interpro; IPR004412; CRN-synt_1.
DR Pfam; PF00658; Acyl_transf; 5.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00107; adh_zinc; 1.
DR Pfam; PF00109; ketoacyl-synt; 5.
DR Pfam; PF02801; ketoacyl-synt_C; 5.
DR Pfam; PF05550; pp-binding; 5.
DR Pfam; PF00975; Thioesterase; 1.
DR TIGRfam; TIGR00128; fadD; 5.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; UNKNOWN_1.
DR PROSITE; PS50075; ACP_DOMAIN; 5.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 5.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 4.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_3.
KW Phosphopantetheine.
SQ SEQUENCE 8817 AA; 933938 MW; 125E1F51F35CDBC1 CRC64;

```

```

Query Match 50.5%; Score 817.5; DB 2; Length 8817;
Best Local Similarity 53.5%; Pred. No. 9,4e-46;
Matches 169; Conservative 42; Mismatches 96; Indels 9; Gaps 2;

QY 1 VFVPPGCGPQWPKMGKELLDASVFRBSVACCAAPAPYDMSVEQVLRSPDAPGLDRV 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7311 VFVPPGCGSQWEGMAISLDDSSVPFPAQLEACCRALAPHEWMLVLRDEGPSPIDRV 7370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 DVVQPTLFAMVIMISLALNMSGVPCAVLGSHSGEIAAAHVSGLSLADAAVYTLMSQA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7371 DVVQPALFAVWVSLALNMSGVGPAAVGHSGCEIAAAVAAALSLIEDNAKIALNSKA 7430
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 QTLTAGTALVSYAATPDELPRIPATPTEDNPARLAAVANGRSTVSGAREVADLVA 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7431 LTTVANGMAAAVELGASDQLGTALPMGD----RLSIAAANSPRATLVSGEPAALDALID 7486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 DLTAQVTRMIPVDVPAHSPLMYAIEERVSGLLPTPRSRIPPHSSVTGRLDPTREL 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7487 SLTPAQVFARKVAVDVASHSAQMDAVODELAAGLANIAPRTCELPLYSTVGTGRLDSEL 7546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 DAAVWYRNNSSTVFEPAPARLLIQGPKTFVEMSPHPVLTMGLOELAPDLGDTTGTADTV 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7547 DGAWYRNLRQTLVLFSSATERRLLDDGHRFVEVSPHVLTLALRETC-----ERSPLDVP 7601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 IMGTLRGGCTLDHFL 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7602 VVGSIRDEGHARLL 7617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 5
09S0R3
ID 09S0R3 PRELIMINARY; PRT; 4881 AA.
AC 09S0R3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 4.
GN AVEA4.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99380548; PubMed=10449723;
RA Ikeda H., Nonomura T., Usami M., Onca T., Omura S.;
RT "Organization of the biosynthetic gene cluster for the polyketide
RT anthelmintic macroide avermectin in Streptomyces avermitilis."
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL; AB032367; BAA84479.1; -.
DR HSSP; P25715; 1MLA.
DR Interpro; IPR001227; Ac transferase.
DR Interpro; IPR000794; ketoacyl-synt.
DR Interpro; IPR003880; Pantne_attach.
DR Interpro; IPR001031; Thioesterase.
DR Pfam; PF00658; Acyl_transf; 3.
DR Pfam; PF00109; ketoacyl-synt; 3.
DR Pfam; PF02801; ketoacyl-synt_C; 3.
DR Pfam; PF00550; pp-binding; 3.
DR Pfam; PF02801; ketoacyl-synt; 1.
DR Pfam; PF00975; Thioesterase; 1.
DR PROSITE; PS50075; ACP_DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
KW Phosphopantetheine; transferase.
SQ SEQUENCE 4881 AA; 510303 MW; F6566C1F01ACD37A CRC64;

Query Match 48.9%; Score 791; DB 2; Length 4881;
Best Local Similarity 52.2%; Pred. No. 2,7e-44;
Matches 164; Conservative 45; Mismatches 97; Indels 8; Gaps 2;

QY 3 VFPGGPPQWPKMGKELLDASVFRBSVACCAAPAPYDMSVEQVLRSPDAPGLDRV 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2139 LFGGCGTHWEGMAQELDGSSVPFPAQMSDCAQLBYLWMSLIDVLRGAAPDAPPLGRVDV 2198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 63 VQPTLFAVMTSLAALMSQGVPCAVLGHSLGEITAAAVSGSLADAAVVTLMGSAQT 122
DB 2199 VQPTLFAVMTSLAALMSQGVPCAVLGHSLGEITAAAVSGSLADAAVVTLMGSAQT 2258
QY 123 TLAAGTGAALVSAATPDELPRIPAWTEEDNPARLAAVAVNGRSTVSGAREAVADVADL 182
DB 2259 ALAAGGAMASVGLPEVKELEPRLATWGD---RLVIAAVNGARSAVSGEPEADVAVEEL 2314
QY 183 TAAOVRTRMTFVDVPAHSPMLTYAIEERVVSGSLPITRPSRIPHSVVTGRLDTRDLDA 242
DB 2315 SHEDVPAHRLMVDWASHSPQVLAQGRLELLELAPRARTGVPSTVTEERIDGTDLDA 2374
QY 243 AAYWRNNSSTVFEPAARLLIQGPKTFVEMSPHVLTMGLQELAPDLGDTGTADTVIM 302
DB 2375 DVTYRNLKQVRRFDALQALVRAGHTVFIECPHAPAVAVGQETLDEMGD---LDSLVV 2430
QY 303 GTLRGGGTLDHFL 316
DB 2431 GSLRRGGGLRRFL 2444

RESULT 6
053490 PRELIMINARY; PRT; 4151 AA.
AC 053490;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Polyketide synthase.
GN PKS12 OR RV2048C OR MT2108 OR MTV018.35C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Baeham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
RA Hornby S., Jacobs K., Kirogh A., McLean J., Moule S., Murphy J.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares K.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.F., Deboy R., Dodson R., Gwinn M., Haft D., Hickey B.,
RA Kolony J.F., Nelson W.C., Nuyam L.A., Ermolova M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Brehet W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL021899; CAI17262.1; -
DR EMBL; AB007061; AAK46387.1; -
DR HSSP; P25715; IMLA.
DR TIGR; MT2108; -
DR Tuberculost; RV2048C; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002085; Adh zn family.
DR InterPro; IPR001899; Gram_pos anchor.
DR InterPro; IPR000794; Ketocycl-synt.
DR InterPro; IPR003880; Ppanne_attach.
DR Pfam; PF00698; Acyl_transf; 2.
DR Pfam; PF00107; adh_zinc; 2.

DR Pfam; PF00109; Ketocycl-synt; 2.
DR Pfam; PF02801; Ketocycl-synt C; 2.
DR Pfam; PF0550; pp-binding; 2.
DR PROSITE; PS50075; ACP DOMAIN; 2.
DR PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
DR PROSITE; PS00343; GRAM_POS ANCHORING; UNKNOWN 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN 2.
KW Phosphopantetheine; Transferase; Complete proteome.
FT CONFLICT 2147 2147 H -> O (IN REF. 2).
FT CONFLICT 2260 2261 DI -> TV (IN REF. 2).
FT CONFLICT 2268 2268 W -> H (IN REF. 2).
FT CONFLICT 2272 2272 S -> A (IN REF. 2).
FT CONFLICT 2279 2280 AF -> PY (IN REF. 2).
FT CONFLICT 2282 2283 AA -> GR (IN REF. 2).
FT CONFLICT 2287 2287 T -> V (IN REF. 2).
FT CONFLICT 2289 2289 F -> W (IN REF. 2).
FT CONFLICT 3004 3004 S -> L (IN REF. 2).
FT CONFLICT 3649 3649 P -> A (IN REF. 2).
SQ SEQUENCE 4151 AA; 431577 MW; 536F644F11D7AD5D CRC64;
Query Match 47.8%; Score 773; DB 16; Length 4151;
Best Local Similarity 51.6%; Pred. No. 3,6e-43;
Matches 163; Conservative 45; Mismatches 102; Indels 6; Gaps 2;
QY 1 VFVFPQGGQPMQMGRELLDASDVRESRACAPAPYDMSVEQVLRDSPAPGLDVRV 60
DB 559 VFVFPQGGQPMQMGRELLDASDVRESRACAPAPYDMSVEQVLRDSPAPGLDVRV 618
QY 61 DVVQPTLFAVMTSLAALMSQGVPCAVLGHSLGEITAAAVSGSLADAAVVTLMGSAQT 120
DB 619 DVVQPTLFAVMTSLAALMSQGVPCAVLGHSLGEITAAAVSGSLADAAVVTLMGSAQT 120
QY 121 QTLAAGTGAALVSAATPDELPRIPAWTEEDNPARLAAVAVNGRSTVSGAREAVADVADL 180
DB 679 QTLAAGTGAALVSAATPDELPRIPAWTEEDNPARLAAVAVNGRSTVSGAREAVADVADL 180
QY 181 DLTAOVRTRMTFVDVPAHSPMLTYAIEERVVSGSLPITRPSRIPHSVVTGRLDTRDLDA 240
DB 735 DLTAOVRTRMTFVDVPAHSPMLTYAIEERVVSGSLPITRPSRIPHSVVTGRLDTRDLDA 240
QY 241 DAAYWRNNSSTVFEPAARLLIQGPKTFVEMSPHVLTMGLQELAPDLGDTGTADTVIM 300
DB 795 DAAYWRNNSSTVFEPAARLLIQGPKTFVEMSPHVLTMGLQELAPDLGDTGTADTVIM 300
QY 301 IMGTLRGGGTLDHFL 316
DB 853 IMGTLRGGGTLDHFL 316
RESULT 7
09LAW4 PRELIMINARY; PRT; 3192 AA.
AC 09LAW4;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE MySB.
GN MYSB.
OS Streptomyces noursei.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1971;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11455;
RX MEDLINE=20334850; PubMed=10873841;
RA Brautaset T., Sekurova O.N., Sletta H., Billingsen T.E., Strom A.R.,
RA Valla S., Zotchev S.B.;
RT "Biosynthesis of the polyene antifungal antibiotic nystatin in
RT Streptomyces noursei ATCC 11455: analysis of the gene cluster and
RT deduction of the biosynthetic pathway."
RL Chem. Biol. 7:395-403(2000).
DR EMBL; AF263912; AAF71775.1; -

DR InterPro: IPR001227; Ac_transferrase.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR003880; Pantone_attach.
DR Pfam: PF00698; Acyl_transf_2.
DR Pfam: PF00109; ketoacyl-synt; 2.
DR Pfam: PF02801; ketoacyl-synt_C_2.
DR Pfam: PF00550; pp-binding; 2_C_2.
DR PROSITE: P850075; ACP_DOMAIN; 2.
DR PROSITE: P800606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE: P800012; PHOSPHOPANTHETINE; UNKNOWN_2.
KW Phosphopantetheine; Transferase.
SEQUENCE 3192 AA; 333916 MW; P88R0864F086DA64 CRC64;

Query Match 47.3%; Score 765.5; DB 2; Length 3192;
Beet Local Similarity 50.6%; Pred. No. 8,3e-43;
Matches 160; Conservative 42; Mismatches 109; Indels 5; Gaps 2;

QY 1 VFVFGGSGPQNGMGRELLDDASDVRESVRACEAFAFYVDMVSEVGRDSPDAPGDRV 60
DB 2114 VFVFGGSGQMGKMRRELLATSEVTRKTIIDCATATAYVDMSLHDVLAGGCDPALLERV 2173
QY 61 DVQVPTLPAVMISLALWRSQGVBCAVLGHSLGEIAAAHYSGSLSDAARVYTLWSQA 120
DB 2174 DVQVQALPAMVYGSLMRSHGVPAVAVGHSGEITAAACVAGSLDAARVVALNSQA 2233
QY 121 OTTLAGTGAIVSAATPEDELPRIAPWTEDDPARLAVAANGPSTVTSGARAVADIVA 180
DB 2234 LPQLSGRGGMVSGVSPVERVTALLPMOE---ALSVAAVNGPSSVWVSGTDLIDLALHT 2289
QY 161 DLTAAQVTRMIPVVPVAPHSPLMVYIERVVSGLPIPRPSRIPIHSSVNGRLDREL 240
DB 2230 ACQEGQVRRKRVSVYASHGRHVEVARELRLARLVAPVPRAPVEVFTVVGDRDDAA 2349
QY 241 DAAVYRWMSSTVREPEARLLIQGQPTVEMSEPHVLTMLGLOELADLDGTTGTAADV 300
DB 2350 DGAVYVYTLRQVVRMEATRALLAAGHVFIEVSHPVLAIPQGTQAVALEATG-GSAY 2408
QY 301 IWGTLRGGQTLDHFL 316
DB 2409 VLGSLRDEGGRFL 2424

RESULT 8

Q9ALM6 PRELIMINARY; PRT; 2595 AA.

AC Q9ALM6; 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Polyketide synthase loading and extender module 1.
GN SPNA.
OS Saccharopolyspora spinosa.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Pseudonocardiales; Pseudonocardiceae;
OC Saccharopolyspora.
OX NCBI_TaxID=60894; [1]
RP SEQUENCE FROM N.A.
RA Weldon C., Matsushita P., Rosebeck P. R. Jr., Broughton M. C.,
RA Turner J., Madduri K., Crawford K. P., Merlo D. J., Baltz R. H.,
RT "Cloning and analysis of the spinosad biosynthetic cluster of
RT Saccharopolyspora spinosa.",
RL Chem. Biol. 8:487-499 (2001).
DR EMBL: AY007554; AAG23264.1; -
DR InterPro: IPR001227; Ac_transferrase.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR003880; Pantone_attach.
DR Pfam: PF00698; Acyl_transf_2.
DR Pfam: PF00106; adh_short; 1.
DR Pfam: PF00109; ketoacyl-synt; 2.
DR Pfam: PF02801; ketoacyl-synt_C_2

Query Match	47.3%	Score 765	DB 2	Length 2595
Best Local Similarity	50.6%	Pred. No. 6.9e-43		
Matches 157	Conservative 51	Mismatches 92	Indels 10	Gaps 3
QY	1	VFEVPGGQPGMGRELLDSDVFRSVCRAEAPPYVDKMSVGYLRDSDPADGDRV	60	
DB	527	VFEVPGGQSGMTGGRQLRLESEFAASMSDCADALPHLDMSLLDVRNAGAAHLDH	586	
QY	61	DVQVPTLFAVWISIAALMRSGCVPCAVLGHSLGEIAAAVHSGSLADAAVVTLSQA	120	
DB	587	DVQVPAFLAVISLAEIMRSGVPAVAVGHSGGEIAAACVAGLSTRDAARVAVASRL	646	
QY	121	QTTIAGGALVSVAATDELLPRTAPMTEDNPAFLAAVANGPSTVSGAREAVDLVA	180	
DB	647	LTLAAGSGAVALSLQHPREAEVQILLPRD----RIGAGVNGPSTVSGDREMAELLA	702	
QY	181	DLTAAGVTRTIPDVAHSPMLVAIEERVSGLLPTPRPSRI PFHSVYGRDLTREL	240	
DB	703	ECADRELKMRIPAEVASHSHIIVVDELGLLPAVPEPRGSLPTVSTTGDLDD-RPM	761	
QY	241	DAAYVYRMSSTVFEPFAPARLLQQGPKTEVMSPHFVLTMGLOELAPDLGTTGTA	300	
DB	762	DADVYRNLRLPVLFEAAVELLRGDAFIEISPHFVLTAINTDETAVRAG-----REV	816	
QY	301	IMGTLRRGQC 310		
DB	817	ALGTLRRGEG 826		
RESULT 9				
Q93HU3	PRELIMINARY	PRT	3939	AA.
AC	Q93HU3			
DT	01-DEC-2001 (TREMBLREL. 19, Created)			
DT	01-DEC-2001 (TREMBLREL. 19, Last sequence update)			
DT	01-MAR-2002 (TREMBLREL. 20, Last annotation update)			
DE	Modular polyketide synthase.			
GN	OLM43.			
OS	Streptomyces avermitilis.			
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
OC	Actinomycetales; Streptomyocineae; Streptomycetaceae; Streptomyces.			
OX	NCBI_Taxid=33903;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21477403; PubMed=11572949;			
RA	Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C. T.,			
RA	Shinoe M., Takahashi Y., Horikawa H., Nakazawa H., Osoboe T.,			
RA	Kikuchi H., Shida T., Sakaki Y., Hattori M.;			
RT	"Genome sequence of an industrial microorganism Streptomyces			
RT	avermitilis": Deducing the ability of producing secondary			
RT	metabolites";			
RU	Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).			
DR	EMBL; AB070940; BAB69194.1;			
DR	InterPro; IPR002106; AATRNA_11gaaeit.			
DR	InterPro; IPR001227; Ac transferase.			
DR	InterPro; IPR002085; adh zn family.			
DR	InterPro; IPR000794; ketocacyl-synt.			
DR	InterPro; IPR003880; pantane_atctch.			
DR	InterPro; IPR002364; QOR_zeta_cystal.			
DR	Pfam; PF00698; Acyl_transfer; 2.			
DR	Pfam; PF00107; adh_zinc; 1.			
DR	Pfam; PF00109; ketocacyl-synt; 2.			
DR	Pfam; PF00801; ketocacyl-synt_C_2.			
DR	Pfam; PF00550; pp-binding; 2.			
DR	PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.			
DR	PROSITE; PS00075; ACP_DOMAIN; 2.			

DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 DR PROSITE; PS01162; QOR_ZETA_CRYSTAL; UNKNOWN_1.
 KW Phosphopantetheine.
 SQ SEQUENCE 3939 AA; 410482 MW; 534AE3B517256503 CRC64;

Query Match 47.2%; Score 763; DB 2; Length 3939;
 Best Local Similarity 50.2%; Pred. No. 1.6e-42;
 Matches 157; Conservative 43; Mismatches 103; Indels 10; Gaps 4;

QY 1 VFVPPGGQPGWPGREGLLDASDVFPRESVACAAAPAYDMSVEQVLRDPPAPGLDRV 60
 DB 572 VFVPPGGQSGQWAGARBLDLSVPFAAKAAIAEFPDWKVLDTLRKPGGAASMERI 631
 QY 61 DVVQPTLFAVMISALALMSQGVPCAVLGSHLGEIAAAHVSGSLADAAVYTLMSQ 119
 DB 632 EVIOPVLFYVWVSLAEIMRSYGVEPDAVSSQGEIAAAHVAGSLTODAAKAVCLMSAL 691
 QY 120 AQTTLAAGTALVSAATPDELPRIPAWTEENPARLVAAVNGRSTVSGAREAVADLV 179
 DB 692 LAETLVGKGAVASVALPADEVERELTRWAD---RLSVAAGVNGPRIVAAGDDALAEFI 747
 QY 180 ADLTAQVTRRMIVDVPAHSPLMVAIBERVVSGLLPTPRPSRIPPHSSVTGRLDTRE 239
 DB 748 AECARDNVRAIRIVATVPTHCAVDPRERLIDLAPVAPRTGTPVSTVTGRLDTAA 807
 QY 240 LDAAYVYRNMSSTVREPPARLLLOQGPRTFVMSPPHVLTMGLQELAPDLGDTTGTADT 299
 DB 808 MDAGYWDNTRAPVLFEPVYRALLABGHSAPVSSAPVLAANGVEQIV---DMAAG-ASG 862
 QY 300 VINGTLRRGGQTL 312
 DB 863 VVVESLRRDEGGL 875

RESULT 10

033956 PRELIMINARY; PRT; 3729 AA.

ID 033956;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE TyLactone synthase modules 4 & 5.
 GN TYLG.
 OS Streptomyces fradiae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1906;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA DeHoff B.S., Sutton K.L., Roestek P.R. Jr.;
 RT "Sequence of Streptomyces fradiae tyLactone synthase gene tyLG.";
 RT Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U78289; AB86506.1; -
 DR InterPro; IPR002106; AATRNA_ligaseII.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh zn family.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Ppanne_attach.
 DR InterPro; IPR002364; QOR_zeta_crystal.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00698; Acyl_transf. 2.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt. 2.
 DR Pfam; PF02801; ketoacyl-synt_C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 DR PROSITE; PS50075; ACP_DOMAIN; 2.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 DR PROSITE; PS01162; QOR_ZETA_CRYSTAL; 1.

DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
 KW Phosphopantetheine; Transferase.
 SQ SEQUENCE 3729 AA; 389375 MW; 19483A5B8027BAC1 CRC64;

Query Match 47.1%; Score 762; DB 2; Length 3729;
 Best Local Similarity 50.9%; Pred. No. 1.7e-42;
 Matches 161; Conservative 46; Mismatches 99; Indels 10; Gaps 4;

QY 1 VFVPPGGQPGWPGREGLLDASDVFPRESVACAAAPAYDMSVEQVLRDPPAPGLDRV 60
 DB 2096 VFVPPGGQSGQWAGARBLDLSVPFAAKAAIAEFPDWKVLDTLRKPGGAASMERI 2155
 QY 61 DVVQPTLFAVMISALALMSQGVPCAVLGSHLGEIAAAHVSGSLADAAVYTLMSQ 120
 DB 2156 DVVQPTLFTVWVSLAEIMRSYGVEPDAVSSQGEIAAAHVAGSLTODAAKAVCLMSAL 2215
 QY 121 QTTTLAAGTALVSAATPDELPRIPAWTEENPARLVAAVNGRSTVSGAREAVADLV 180
 DB 2216 WTLTAGKGAVASVALPADEVERELTRWAD---RLSVAAGVNGPRIVAAGDDALAEFI 2271
 QY 181 ADLTAQVTRRMIVDVPAHSPLMVAIBERVVSGLLPTPRPSRIPPHSSVTGRLDTRE 239
 DB 2272 ELTAEGIRAKPRTGVDTAGHSAQVGGKEHLEFVLAPVSRSSDIPYISTVTAAPLDTRE 2331
 QY 240 LDAAYVYRNMSSTVREPPARLLLOQGPRTFVMSPPHVLTMGLQELAPDLGDTTGTADT 299
 DB 2332 LDAAGYWDNTRAPVLFEPVYRALLABGHSAPVSSAPVLAANGVEQIV---DMAAG-ASG 2387
 QY 300 VINGTLRRGGQTL 315
 DB 2388 -VWHTLRROKSAKDP 2402

RESULT 11

093H18 PRELIMINARY; PRT; 3970 AA.

ID 093H18;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Modular polyketide synthase.
 GN OLM44.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Iehikawa J., Hanamoto A., Takahashi C.,
 RA Shinohe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: Deducing the ability of producing secondary
 metabolites.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL; AB070940; BAB6919.1; -
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh_zn_family.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Ppanne_attach.
 DR InterPro; IPR002364; QOR_zeta_crystal.
 DR Pfam; PF00698; Acyl_transf. 2.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt. 2.
 DR Pfam; PF02801; ketoacyl-synt_C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PS50075; ACP_DOMAIN; 2.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 DR PROSITE; PS01162; QOR_ZETA_CRYSTAL; UNKNOWN_1.
 KW Phosphopantetheine.
 SQ SEQUENCE 3970 AA; 412657 MW; A93E583FAFA68C47 CRC64;

Db 3851 VFVPPGSGSQWAGCAQLLDES PFVFAERIEACAAALAEFTDMLNLDVLRGAEAPTLERV 3910
QY 61 DVVQPTLFAMVIMSLAALMRSQGVPCAVLGHSLGEIAAAHVSGGLSLADARVVTLMQA 120
Db 3911 DVVQPSAFVAVWSLAAVWRAGVBPDAVVGHSQGEIAAAVSGALSLRDGARVVTLMQA 3970
QY 121 -QTTLAGTGAIVSAATPEDELLPRIAPWTEEDNPARLAAVAVNGRSTVSGAREAVADLV 179
Db 3971 IGRSLAGRGGMWSVALPVAEVEARL---EAFEGRVSAVAAENGPRSSVAGPEPALDELH 4026
QY 180 ADLTAQVTRFMIVDVDPVAPHSPLMTVAIEERVVSGLLPTTPRSRIAPHSSVTGRLDTR 239
Db 4027 AQLTAEEIRARAVADVASHSPHVEDHDELLELAEVAPRTSEIPFFSTYTGMDLDTTV 4086
QY 240 LDAAYWRNMSSTVRFEPAARLLQOGPKTFVMSPPHVTMGLOELAPDLGDTTGTADT 299
Db 4087 MDAGYTRSLRGRVLPADAVADLTAADHRATIEVSHPVLMASVQDMIDDAG-VAGVAS- 4144
QY 300 VINGTLRRCGGTLDHFL 316
Db 4145 ---GTLRRDNGGLDRFL 4158

RESULT 14

Q9EMAI PRELIMINARY; PRT; 9507 AA.
ID O9EMAI
AC O9EMAI
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE PIMS2 protein.
GN PIMS2.
OS Streptomyces natalensis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=68242;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20547809; PubMed=11094342;
RA Aparicio J.F., Fouces R., Mendes M.V., Oliveira N., Martin J.F.;
RT "A complex multienzyme system encoded by five polypeptide synthase
genes is involved in the biosynthesis of the 26-membered polyene
macrolide pimarsin in Streptomyces natalensis.";
RT Chem. Biol. 7:895-905(2000).
RL EMBL: AJ278573; CAC30921.1; -.
DR HSSP: P25715; IMLA
DR InterPro: IPR001227; Ac transferase.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR004410; FAD.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR000794; ketoacyl-synt.
DR InterPro: IPR001005; Myb_DNA_binding.
DR InterPro: IPR003880; Ppantne_attach.
DR Pfam: PF00698; Acyl_transf. 6.
DR Pfam: PF00106; adh_short. 3.
DR Pfam: PF02801; ketoacyl-synt. 6.
DR Pfam: PF00550; pp-binding. 6.
DR TIGRfam: TIGR00128; fad. 6.
DR PROSITE: PSS0075; ACP_DOMAIN. 6.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE. 6.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE. 5.
KW Phosphopantetheine.
SQ SEQUENCE 9507 AA; 994228 MW; 57BD80C8AD37E89F CRC64;

Query Match 46.6%; Score 754.5; DB 2; Length 9507;

Best Local Similarity 51.4%; Pred. No. 1.8e-41;

Matches 163; Conservative 43; Mismatches 100; Indels 11; Gaps 4;

QY 1 VFVPPGSGSQWAGCAQLLDES PFVFAERIEACAAALAEFTDMLNLDVLRGAEAPTLERV 60

Db 3859 VFVPPGSGSQWAGCAQLLDES PFVFAERIEACAAALAEFTDMLNLDVLRGAEAPTLERV 3918
QY 61 DVVQPTLFAMVIMSLAALMRSQGVPCAVLGHSLGEIAAAHVSGGLSLADARVVTLMQA 120
Db 3919 DVVQPSAFVAVWSLAAVWRAGVBPDAVVGHSQGEIAAAVSGALSLRDGARVVTLMQA 3978
QY 121 -QTTLAGTGAIVSAATPEDELLPRIAPWTEEDNPARLAAVAVNGRSTVSGAREAVADLV 179
Db 3979 IGRSLAGRGGMWSVALPVAEVEARL---EAFEGRVSAVAAENGPRSSVAGPEPALDELH 4033
QY 180 ADLTAQVTRFMIVDVDPVAPHSPLMTVAIEERVVSGLLPTTPRSRIAPHSSVTGRLDTR 239
Db 4034 AQLTAEEIRARAVADVASHSPHVEDHDELLELAEVAPRTSEIPFFSTYTGMDLDTTV 4093
QY 240 LDAAYWRNMSSTVRFEPAARLLQOGPKTFVMSPPHVTMGLOELAPDLGDTTGTADT 299
Db 4094 MDAGYTRSLRGRVLPADAVADLTAADHRATIEVSHPVLMASVQDMIDDAG-VAGVAS- 4148
QY 300 VINGTLRRCGGTLDHFL 316
Db 4149 AVIGTLRRDNGGLDRFL 4165

RESULT 15

Q9LAX3 PRELIMINARY; PRT; 9477 AA.
ID O9LAX3
AC O9LAX3
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Nysl.
GN NYSI.
OS Streptomyces noursei.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1971;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 11455;
RC MEDLINE=20334850; PubMed=10873841;
RA Brautaset T., Sekurova O.N., Sletta H., Ellingsen T.E., Strom A.R.,
RA Valla S., Zolchev S.B.;
RT "Biosynthesis of the polyene antifungal antibiotic nystatin in
Streptomyces noursei ATCC 11455: analysis of the gene cluster and
RT deduction of the biosynthetic pathway.";
RT Chem. Biol. 7:395-403(2000).
RL EMBL: AF263912; AF71766.1; -.
DR HSSP: AF263912; AAF71766.1; -.
DR HSSP: P81989; IQES.
DR InterPro: IPR001227; Ac transferase.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR004410; FAD.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR000794; ketoacyl-synt.
DR InterPro: IPR001005; Myb_DNA_binding.
DR InterPro: IPR003880; Ppantne_attach.
DR Pfam: PF00698; Acyl_transf. 6.
DR Pfam: PF00106; adh_short. 3.
DR Pfam: PF02801; ketoacyl-synt. 6.
DR Pfam: PF00550; pp-binding. 6.
DR TIGRfam: TIGR00128; fad. 6.
DR PROSITE: PSS0075; ACP_DOMAIN. 6.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE. 6.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE. 6.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 9477 AA; 988132 MW; 0D2BCA5D7B2E5483 CRC64;

Query Match 46.5%; Score 753; DB 2; Length 9477;

Best Local Similarity 53.0%; Pred. No. 2.3e-41;

Matches 168; Conservative 34; Mismatches 105; Indels 10; Gaps 3;

```
Qy 1 VFVPCGGPQWPKNGRELDASDVFRBSVRACEAFAFVVDMSVEQVLRDSPDAPGLDRV 60
Db 3895 VFVPCGGSQWVGMSQLDDESAVFAERIAECAAAALAEPTDWSLVDTLRGVGAPSLERV 3954
Qy 61 DVVQPTLFAVMISIAALMBSQGVPCAVLGHSLGEIAAAHVSGGLSLADAARVYTLMSQA 120
Db 3955 DVVQPAFAVMVSLAALMBSRGVLPDAVGHSGEIAAAVVSGLSLRDGARVVALRSQA 4014
Qy 121 -QTLAAGTALVSVAAATPDELLPRIAPWTEDNPARLVAAVNGPRSTVSGARBAVADLV 179
Db 4015 IGRALAGRGGMVSVALSVDLERLVEF---EGRVSAAVNGPRSVVAGPEPALDALH 4070
Qy 180 ADLTAAGVTRMIIPVDVPAHSPLMTAIEBRSVSGLPITPRPSRIIPHSVTGRLDTRE 239
Db 4071 ARLTAADDIRARIAVDYASHSHQVEDHELEVLAEIAPRTSEVPFSTVTGMDLDTAR 4130
Qy 240 LDAAYVYRNMSSTVREPPAARLLLOQGPKTFFVEMSPHPUVTMGLQELAPDLGDTTGTADT 299
Db 4131 MDAGYWFRRNLGRVRFADAVADLLAAEYRAFEVSSHPLSMVQEAIDEAG----VPA 4185
Qy 300 VIMGTLRGGQTLDHFL 316
Db 4186 VAAGTILRRDGGGTDRFL 4202
```

Search completed: June 17, 2003, 13:07:51
Job time : 15.6048 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 17, 2003, 12:52:22 ; Search time 3.44608 Seconds
(without alignments)
3803.313 Million cell updates/sec

Title: US-09-914-286-3_COPY_29_344
Perfect score: 1618
Sequence: 1 VFPVPGGSPQWPGWGRELD.....ADTVIMGTLRGGTLDHFL 316

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	776.5	48.0	3567 1	ERY2_SACER
2	722	44.6	3491 1	ERY1_SACER
3	707	43.7	3519 1	OL56_STRAT
4	689	42.6	3172 1	ERY3_SACER
5	463.5	28.6	2110 1	MCAS_MYCBO
6	410	25.3	1538 1	PPSB_MYCTU
7	398	24.6	1876 1	PPSA_MYCTU
8	366.5	12.7	1774 1	MSAS_PENPA
9	265	16.4	317 1	FABD_BACSU
10	258	15.9	1986 1	WA_EMENI
11	255	15.8	308 1	FABD_SALTY
12	252	15.6	308 1	FABD_ECOLI
13	247	15.3	2109 1	PKS1_ASPPA
14	243.5	15.0	2181 1	STCA_EMENI
15	217	13.4	293 1	FABD_SYNY3
16	213.5	13.2	312 1	FABD_HAEIN
17	205.5	12.7	2505 1	PAS_BAT
18	187	11.6	2511 1	PAS_CHICK
19	167.5	10.4	2504 1	PAS_HUMAN
20	125.5	7.8	2037 1	FAS1_HUMAN
21	123.5	7.6	463 1	Y102_MYCTU
22	123	7.6	2073 1	FAS1_SCHPO
23	109	6.7	1914 1	STCK_EMENI
24	108.5	6.7	920 1	PARC_SYNY3
25	107	6.6	2035 1	HFCL_HUMAN
26	104	6.4	426 1	G6PI_HUMAN
27	103	6.4	547 1	MUTL_DEIRA
28	102.5	6.3	343 1	GPDA_RHILQ
29	101	6.2	494 1	MURC_MYCTU
30	101	6.2	2090 1	HFCL_MESAU
31	100.5	6.2	509 1	YD30_MYCTU
32	99	6.1	375 1	EPC2_RALISO
33	97	6.0	852 1	SYO_DEIRA

34	97	6.0	2051 1	FAS1_YEAST	P07149 s fatty aci
35	96.5	6.0	302 1	FABD_MYCTU	Q10501 mycobacteri
36	96.5	6.0	304 1	LDH_DEIRA	P50923 deinnococcus
37	96	5.9	1074 1	CARB_SALTI	O82917 salmonella
38	96	5.9	1074 1	CARB_SALTY	P14846 salmonella
39	96	5.9	1108 1	CYGE_RAT	P51840 rattus norv
40	96	5.9	1286 1	IRBP_BOVIN	Q05562 streptomyce
41	95.5	5.9	225 1	GRPE_STRCO	Q10813 mycobacteri
42	95.5	5.9	408 1	YS92_MYCTU	Q02846 homo sapien
43	95.5	5.9	1103 1	CYGD_HUMAN	P36748 human papil
44	94.5	5.8	518 1	VL2_HPV12	P58600 ralsionia s
45	94	5.8	375 1	EPC1_RALISO	

ALIGNMENTS

```

RESULT 1
ID ERY2_SACER STANDARD; PRT; 3567 AA.
AC 003132; 054096;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Erythronolide synthase, modules 3 and 4 (BC 2.3.1.94) (ORF 2) (6-
DE deoxyerythronolide B synthase II) (DEBS 2).
GN ERYA
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Actinobacteria; Actinobacteridae (class); Actinobacteridae;
OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae;
OC Saccharopolyspora.
OX NCBI_TaxID=1836;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91220065; PubMed=2024119;
RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.,
RT "Modular organization of genes required for complex polyketide
RT biosynthesis."
RL Science 252:675-679 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=NRRL 2338;
RA Bevil D.J., Cortes J., Haydock S.F., Leadlay P.F.;
RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea.
RT Cloning of the structural gene, sequence analysis and inferred domain
RT structure of the multifunctional enzyme."
RL Eur. J. Biochem. 204:39-49 (1992).
CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
CC deoxyerythronolide B.
CC -1- COFACTOR: NADP. CONTAINS 2 COVALENTLY BOUND PHOSPHOTRANSFERINS.
CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
CC BIOSYNTHESIS.
CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH
CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3
CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
CC SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS
CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6,
CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
CC RESPECTIVELY.
CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES: ACYLTRANSFERRASE (AT),
CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER
CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETO REDUCTASE (KR),
CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
CC OF THE FULL-LENGTH CHAIN.
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M63677; AAA26494.1; -
 DR EMBL; X62569; CAA44448.1; -
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh zn family.
 DR InterPro; IPR000794; Ketocacyl-synt.
 DR InterPro; IPR003880; Pantane_attach.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketocacyl-synt; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl transf; 2.
 DR Pfam; PF02801; ketocacyl-synt C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
 DR PROSITE; PS00606; B-KETOACYL SYNTHASE; 2.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 KM Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KM Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1484 MODULE 3.
 FT DOMAIN 1485 3567 MODULE 4.
 FT DOMAIN 27 488 BETA-KETOACYL SYNTHASE 1.
 FT DOMAIN 559 884 ACYLTRANSFERASE (AT) 1.
 FT DOMAIN 1130 1301 BETA-KETOACYL REDUCTASE 1 (POSSIBLY NON-FUNCTIONAL).
 FT DOMAIN 1397 1467 ACYL CARRIER (ACP) 1.
 FT DOMAIN 1485 1943 BETA-KETOACYL SYNTHASE 2.
 FT DOMAIN 2013 2336 ACYLTRANSFERASE (AT) 2.
 FT DOMAIN 2283 3066 DEHYDRATASE/ENOYL REDUCTASE (DH/ER).
 FT DOMAIN 3139 3322 BETA-KETOACYL REDUCTASE 2.
 FT DOMAIN 3415 3485 ACYL CARRIER (ACP) 2.
 FT ACT_SITE 202 THIOESTER BOND.
 FT ACT_SITE 651 ACYL-ENZYME INTERMEDIATE.
 FT BINDING 1430 1430 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT ACT_SITE 1661 1661 THIOESTER BOND.
 FT NP_BIND 2115 2115 ACYL-ENZYME INTERMEDIATE.
 FT NP_BIND 2961 2978 NADP (KR).
 FT NP_BIND 3142 3157 NADP (KR).
 FT BINDING 3448 3448 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT CONFLICT 438 438 R -> A (IN REF. 2).
 FT CONFLICT 480 480 T -> S (IN REF. 2).
 FT CONFLICT 1241 1241 L -> F (IN REF. 2).
 FT CONFLICT 2664 2664 G -> V (IN REF. 2).
 SQ SEQUENCE 3567 AA; 374413 MW; EB6284F4738AA0C0 CRC64;

Query Match 48.0%; Score 776.5; DB 1; Length 3567;
 Best Local Similarity 50.0%; Pred. No. 2.3e-48;
 Matches 158; Conservative 44; Mismatches 107; Indels 7; Gaps 2;

QY 1 VFVPPGGGQWPGMGRELLDASDVFRSVCRAAPAPYVWVSVQVLRDSDPAQIDRV 60
 DB 560 VFLEPGGQSGWAGAEHLSSPVFAGKIRACDSMAWMDKWSVLRQAPGAPGIDRV 619
 QY 61 DVVQPTLFAVWISLAALMRSGVEPCAVLGHSLGEIAAAHVSGGLSADARVVTLSQA 120
 DB 620 DVVQPTLFAVWISLAALMRSGVEPCAVLGHSLGEIAAAHVSGGLSADARVVTLSQA 120
 QY 121 QTTIAGCALVSVAATPDELPRILAPWTEPNPAFLAANAANGPSTVSGAREVADIVA 180
 DB 680 MRSLSGGGMAAVALGEAAVVERLRPMOD---RLSVAANAANGPSTVSGAREVADIVA 180
 QY 181 DLTAAGVTRMIPVDVPAHSPMLKMAIERVVSGLLPTPRSRIRPHSSVVGSLDREL 240
 DB 736 DCAABEGIRVDIDVYASHSPQIRVRELETTGGDIAPRAKAVTFHSVSGRMDGTEL 795
 QY 241 DAAVYVNMSTVFEPPARALLQGGPTFEVMSPHVLTGGLQELAPDLGDTTGTADTV 300
 DB 796 DAAVYVNMSTVFEPPARALLQGGPTFEVMSPHVLTGGLQELAPDLGDTTGTADTV 300
 QY 301 IMGLTRGGQGLDHLFL 316
 DB 853 VVGSIRHDDGDLASFL 868

RESULT 2

ERYL_SACER STANDARD; PRT; 3491 AA.
 ID ERYL_SACER
 AC Q03131;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Erythronolide synthase, modules 1 and 2 (EC 2.3.1.94) (ORF 1) (6-deoxyerythronolide B synthase I) (DEBS I).
 GN ERYA.
 OS Saccharopolyspora erythraea (Streptomyces erythraeus).
 OC Bacteria; Actinobacteria; Actinobacteriales (class); Actinobacteridae; Actinomycetiales; Pseudonocardineae; Pseudonocardaceae;
 OC Saccharopolyspora.
 CX NCBI_TaxID=1836;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220065; Pubmed=2024119;
 RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
 RT "Modular organization of genes required for complex polyketide biosynthesis";
 RL Science 252:675-679 (1991).
 RN [2]
 RP SEQUENCE OF 3474-3491 FROM N.A.
 RX MEDLINE=93231529; Pubmed=8386127;
 RA Donadio S., Staver M.J.;
 RT "18136, an insertion element in the erythromycin gene cluster of Saccharopolyspora erythraea";
 RL Gene 126:147-151 (1993).
 CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-deoxyerythronolide B.
 CC -1- COFACTOR: NADP. CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTETHEINES.
 CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
 CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
 CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR) DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.
 CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
 CC -1- SIMILARITY: CONTAINS 3 ACYL CARRIER DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 ACYL CARRIER DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR	PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
DR	PROSITE; PS00755; ACP DOMAIN; 3.
KM	Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
KM	phosphopantetheine; Multifunctional enzyme.
FT	DOMAIN 1 1972
FT	DOMAIN 1 1979 3491
FT	DOMAIN 1 375
FT	DOMAIN 414 484
FT	DOMAIN 503 961
FT	DOMAIN 1030 1356
FT	DOMAIN 1611 1794
FT	DOMAIN 1888 1958
FT	DOMAIN 1979 2441
FT	DOMAIN 2507 2854
FT	DOMAIN 3055 3237
FT	DOMAIN 3334 3404
FT	ACT_SITE 145 145
FT	BINDING 447 447
FT	ACT_SITE 677 677
FT	ACT_SITE 1128 1128
FT	NP_BIND 1614 1660
FT	BINDING 1921 1921
FT	ACT_SITE 2148 2148
FT	ACT_SITE 2598 2598
FT	NP_BIND 3058 3104
FT	BINDING 3367 3367
SO	SEQUENCE 3491 AA; 365022 MW; 682BFC32C9FA64 CRC64; PHOSHOPANTHETHEINE (BY SIMILARITY).

Query Match 44.6%; Score 722; DB 1; Length 3491;
 Best Local Similarity 50.3%; Pred. No. 2.1e-44;
 Matches 159; Conservative 42; Mismatches 103; Indels 12; Gaps 5;

Qy	114	VFVFGGSGFPMCEMGRELDIASVFEFSSVACLAAPAPVMDVSQVIVLRSDVLAHGUDRV	60
Db	57	VFVFGGSGAOMAMAGELLGSESRVFPAAMDACARAFEPMTDTTLAQVL--DSPEOS--RRV	113
Qy	61	DVVOPTFLFAMVIGLALIMRSQGVPCAVVGHISGELIAAHVCGSLDADARVYTLMSQA	120
Db	114	EVQVPAIFAVQTSIALIMRSFGVTPPAVVGHSIGELIAAHVCGAAGADAPARAALMSRE	173
Qy	121	QTTTLAAGTALVSAATPDELLPRIAPWTEEDNPARIILVAAVNGPRSTVVGAREAVADLVA	180
Db	174	MIPLVGVGDMAAVALASADEIEPIIAMDD----VLAGVNGPRSVLLTGSPEPVARVQ	225
Qy	181	DLTAAGVRTMIVVDVPAHSPLMYALAEERVSGLPIPRPRIPFHSVYSGRLDREL	240
Db	230	ELSBEGRAOVIVNSMAHSAQVDDIABGKRSLAMFAPGSGSEVPFVYSLTGAVDREL	289
Qy	241	DAAYVYRNMSSTVRFEPARLLAQGPKTIVEMSPHPVLTMLGELADLPDPTTGADTV	300
Db	290	VADYWRRSFPLPVRFDEAIRSALEVGPFFVEASPHPVLAALQOTL----DAEGSSAAV	345
Qy	301	IMGLTRGCGTLDHFL	316
Db	346	V-FTLGRGGGGMRRFL	360

RESULT 3	OL56_STRAT	STANDARD;	PRT;	3519 AA.
ID	OL56_STRAT			
AC	Q07017			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Oleandomycin polyketide synthase, modules 5 and 6.			
GN	ORPB.			
OS	Streptomyces antibioticus.			
OC	Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae			
OC	Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.			
OX	NCBI_TaxID=1890;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94150470; PubMed=8107683;			

```

RA Swan D.G., Rodriguez A.M., Vilches C., Mendez C., Salas J.A.;
RT "Characterisation of a Streptomyces antibiotic gene encoding a type
RL I polyketide synthase which has an unusual coding sequence.";
CC Mol. Gen. Genet. 242:358-362(1994).
CC -1- FUNCTION: MAY BE INVOLVED IN THE BIOSYNTHESIS OF THE OLEANDOMYCIN
CC LACTONE RING.
CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; L09654; AAA19695.1; -.
CC DR HSSP; P25715; IMLA.
CC DR InterPro; IPR001227; Ac transferase.
CC DR InterPro; IPR0000794; Ketoacyl-synt.
CC DR InterPro; IPR003880; Pantene_atatch.
CC DR InterPro; IPR01031; Thioesterase.
CC DR Pfam; PF00109; ketoacyl-synt; 2.
CC DR Pfam; PF00550; pp-binding; 2.
CC DR Pfam; PF00698; Acyl transf; 2.
CC DR Pfam; PF00975; Thioesterase; 1.
CC DR Pfam; PF02801; ketoacyl-synt C; 2.
CC DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
CC DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
CC DR PROSITE; PS50075; ACP DOMAIN; 2.
CC KW Transferase; Acyltransferase; Antibiotic biosynthesis; NADP;
CC KW Phosphopantetheine; Multifunctional enzyme; Repeat.
CC FT DOMAIN 1 ?
CC FT DOMAIN ? 3519
CC FT DOMAIN 32 501
CC FT DOMAIN 569 890
CC FT DOMAIN 1200 1382
CC FT DOMAIN 1487 1561
CC FT DOMAIN 1686 2156
CC FT DOMAIN 2220 2541
CC FT DOMAIN 2856 3038
CC FT DOMAIN 3141 3215
CC FT ACT_SITE ?
CC FT BINDING ?
CC FT BINDING ? ?
CC FT DOMAIN 3270 3519
CC FT ACT_SITE 210 210
CC FT ACT_SITE 660 660
CC FT NP_BIND 1203 1249
CC FT BINDING 1524 1524
CC FT ACT_SITE 1859 1859
CC FT ACT_SITE 2311 2311
CC FT NP_BIND 2859 2905
CC FT BINDING 3178 3178
CC FT SEQUENCE 3519 AA; 368561 MW; 41AE78AAAB61F86 CRC64;

```

	Query Match	Similarity	Score	DB 1	length
Beet	Local	49.4%	Pred. No. 2.6e-43		
Matches	157	Conservative	41	Mismatches	106
				Indels	14
				Gaps	4

```

Db      2394 LAGCEGCVARARVPVYASHAQMDDQLRDELBALADITQDSSVFPFSTVTDWLDTT 2453
Cc      239 ELDAAVYRNMSSTVRFEPARLLLOCPKTFVEMSPVUTMLOGLAPLGGTTGAD 298
Cc      2454 ALDAGYFNTNRETFVDEAEGLVAQGMGAFVBCSPHVLVPGLIOTLALDO----N 2508
Cc      299 TVINGTLRRGQGLDHL 316
Db      2509 AAVIGSLRRDEGGDRL 2526

RESULT 4
ID      ERY3 SACER STANDARD; PRT; 3172 AA.
AC      001133: 054097; 099270;
DT      01-OCT-1993 (Rel. 27, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Erythronolide synthase, modules 5 and 6 (EC 2.3.1.94) (ORF 3) (6-
GN      deoxyerythronolide B synthase III) (DEBS 3).
OS      Saccharopolypora erythraea (Streptomyces erythraeus).
OC      Bacteria; Actinobacteria; Actinobacteriota (class); Actinobacteridae;
OC      Actinomycetales; Pseudonocardineae; Pseudonocardiaceae;
OC      Saccharopolypora.
OX      NCBI_TaxID=1836;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=NRRL 2338;
RX      MEDLINE=91043075; PubMed=2234082;
RA      Cortes J., Haydock S.F., Roberts G.A., Bevilitt D.J., Leadlay P.F.;
RT      "An unusually large multifunctional polypeptide in the erythromycin-
RL      producing polyketide synthase of Saccharopolypora erythraea.";
RN      Nature 348:176-178(1990).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91220065; PubMed=2024119;
RA      Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT      "Modular organization of genes required for complex polyketide
RL      biosynthesis.";
RN      Science 252:675-679(1991).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      STRAIN=NRRL 2338;
RX      MEDLINE=92155230; PubMed=1740151;
RA      Bevilitt D.J., Cortes J., Haydock S.F., Leadlay P.F.;
RT      "6-Deoxyerythronolide-B synthase 2 from Saccharopolypora erythraea.
RT      Cloning of the structural gene, sequence analysis and inferred domain
RL      structure of the multifunctional enzyme.";
RN      Eur. J. Biochem. 204:39-49(1992).
Cc      -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
Cc      deoxyerythronolide B.
Cc      -1- COFACTOR: NADP; CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETHEINES.
Cc      -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
Cc      BIOSYNTHESIS.
Cc      -1- MISCELLANEOUS: IN EACH ORF OF ERY3 TWO MODULES ARE PRESENT EACH
Cc      ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERY3 SHOWING 3
Cc      ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
Cc      SYNTHASE PARTICIPATES IN ONE OF THE SIX FAS-LIKE ELONGATION STEPS
Cc      REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
Cc      AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
Cc      RESPECTIVELY.
Cc      -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),
Cc      BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER
Cc      PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
Cc      DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
Cc      BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
Cc      OF THE FULL-LENGTH CHAIN.
Cc      -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
Cc      -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
Cc      -----
Cc      This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc      between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc      the European Bioinformatics Institute. There are no restrictions on its
Cc      use by non-profit institutions as long as its content is in no way
Cc      modified and this statement is not removed. Usage by and for commercial
Cc      entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc      or send an email to license@isb-sib.ch).
Cc      -----
Cc      EMBL; X56107; CA39583.1; -
Cc      DR EMBL; M63677; AAA26495.1; -
Cc      DR EMBL; X62569; CAA44449.1; -
Cc      DR HSP; P00101; ICCH.
Cc      DR InterPro; IPR002198; ADH_short.
Cc      DR InterPro; IPR001227; Ac transferase.
Cc      DR InterPro; IPR000794; Ketoacyl-synt.
Cc      DR InterPro; IPR003880; Pantane attach.
Cc      DR InterPro; IPR001031; Thioesterase.
Cc      DR Pfam; PF00106; adh_short; 1.
Cc      DR Pfam; PF00109; ketoacyl-synt; 2.
Cc      DR Pfam; PF00550; pp-binding; 2.
Cc      DR Pfam; PF00698; Acyl transf; 2.
Cc      DR Pfam; PF00975; Thioesterase; 1.
Cc      DR Pfam; PF02801; ketoacyl-synt C; 2.
Cc      DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
Cc      DR PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
Cc      DR PROSITE; PS50075; ACP DOMAIN; 2.
Cc      KM Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
Cc      KW Phosphopantetheine; Multifunctional enzyme.
Cc      FT DOMAIN 1 1484
Cc      FT 1 1485
Cc      FT DOMAIN 2 3172
Cc      FT 3173 484
Cc      FT DOMAIN 3 7
Cc      FT 8 878
Cc      FT DOMAIN 4 1116
Cc      FT 1117 1298
Cc      FT DOMAIN 5 1394
Cc      FT 1395 1464
Cc      FT DOMAIN 6 1488
Cc      FT 1489 1954
Cc      FT DOMAIN 7 2021
Cc      FT 2022 2335
Cc      FT DOMAIN 8 2555
Cc      FT 2556 2735
Cc      FT DOMAIN 9 2821
Cc      FT 2822 2891
Cc      FT DOMAIN 10 2926
Cc      FT 2927 3172
Cc      FT ACT_SITE 199
Cc      FT ACT_SITE 643
Cc      FT NP_BIND 1118
Cc      FT 1119 1164
Cc      FT BINDING 1427
Cc      FT 1428 1427
Cc      FT ACT_SITE 1661
Cc      FT 1662 2112
Cc      FT NP_BIND 2557
Cc      FT 2558 2605
Cc      FT BINDING 2854
Cc      FT 2855 2854
Cc      FT CONFLICT 231
Cc      FT 232 231
Cc      FT CONFLICT 240
Cc      FT 241 240
Cc      FT CONFLICT 289
Cc      FT 290 493
Cc      FT CONFLICT 493
Cc      FT 494 517
Cc      510 510
Cc      511 513
Cc      512 525
Cc      513 536
Cc      514 547
Cc      515 551
Cc      516 553
Cc      517 553
Cc      518 673
Cc      519 716
Cc      520 734
Cc      521 896
Cc      522 896
Cc      523 896
Cc      524 988
Cc      525 994
Cc      526 1108
Cc      527 1116
Cc      528 1124
Cc      529 1126
Cc      530 1132
Cc      531 1132
Cc      532 1192
Cc      533 1192
Cc      534 1194
Cc      535 1194
Cc      536 1277
Cc      537 1278
Cc      538 1385
Cc      539 1390
Cc      540 1485
Cc      541 1485
Cc      542 1518
Cc      543 1518
Cc      MODULE 5.
Cc      MODULE 6.
Cc      BETA-KETOACYL SYNTHASE 1.
Cc      ACYLTRANSFERASE (AT) 1.
Cc      BETA-KETOACYL REDUCTASE 1.
Cc      ACTYL CARRIER (ACP) 1.
Cc      BETA-KETOACYL SYNTHASE 2.
Cc      ACYLTRANSFERASE (AT) 2.
Cc      BETA-KETOACYL REDUCTASE 2.
Cc      ACTYL CARRIER (ACP) 2.
Cc      ACYL-ENZYME INTERMEDIATE.
Cc      NADP.
Cc      PHOSPHOPANTHETHEINE (BY SIMILARITY).
Cc      THIOESTER BOND.
Cc      ACYL-ENZYME INTERMEDIATE.
Cc      NADP.
Cc      PHOSPHOPANTHETHEINE (BY SIMILARITY).
Cc      MISSING (IN REF. 2).
Cc      A -> R (IN REF. 2).
Cc      P -> R (IN REF. 2).
Cc      PEPRNSLRDGLATRASMEHRA -> ASRGTRCATPVS
Cc      RMEPAAPMEQ (IN REF. 1).
Cc      A -> R (IN REF. 2).
Cc      M -> W (IN REF. 2).
Cc      E -> D (IN REF. 2).
Cc      R -> G (IN REF. 2).
Cc      GPNSP -> ARTR (IN REF. 2).
Cc      R -> A (IN REF. 2).
Cc      MISSING (IN REF. 2).
Cc      AHR -> GIT (IN REF. 2).
Cc      R -> ROR (IN REF. 2).
Cc      R -> RELPYROROR (IN REF. 1).
Cc      GUAUVH -> VALSLRD (IN REF. 2).
Cc      RHPLEPLA -> ARTRMEPR (IN REF. 2).
Cc      MISSING (IN REF. 2).
Cc      L -> V (IN REF. 2).
Cc      A -> R (IN REF. 2).
Cc      MISSING (IN REF. 2).
Cc      AA -> RR (IN REF. 2).
Cc      LQDGR -> STVR (IN REF. 2).
Cc      MISSING (IN REF. 2).
Cc      G -> R (IN REF. 2).

```

FT	CONFLICT	1601	1601	V -> L (IN REF. 2).	
FT	CONFLICT	1724	1725	LP -> PA (IN REF. 2).	
FT	CONFLICT	1732	1732	O -> L (IN REF. 2).	
FT	CONFLICT	1739	1743	CPAG -> ARRA (IN REF. 2).	
FT	CONFLICT	1762	1762	T -> S (IN REF. 2).	
FT	CONFLICT	2252	2252	D -> DGAD (IN REF. 2).	
FT	CONFLICT	2275	2277	OSP -> AVA (IN REF. 2).	
FT	CONFLICT	2408	2408	G -> GR (IN REF. 2).	
FT	CONFLICT	2420	2421	LA -> S (IN REF. 2).	
FT	CONFLICT	2443	2444	NA -> TH (IN REF. 2).	
FT	CONFLICT	2596	2596	A -> G (IN REF. 2).	
FT	CONFLICT	2609	2609	P -> A (IN REF. 2).	
FT	CONFLICT	2715	2722	RRABERAA -> AVKAVRR (IN REF. 1).	
FT	CONFLICT	2754	2754	D -> E (IN REF. 2).	
SQ	SEQUENCE	3172 AA;	331474 MW;	DBBD5094E77DDDSF CRC64;	
Query Match					42.6%;
Beet Local Similarity					46.8%;
Matches 148; Conservative					52; Mismatches 10; Indels 14; Gaps 5

Qy	3	FFPGGGPPMPGKRELLDSDVFRSSVRCFAAFYVYMSEYQVLRDSPDAFGLDRVY	62
Db	557	VFPFGAGMOCMARDLRESQVFPADSIKCEKALAPHVMSLTDL---SGARPLDRVY	613
Qy	63	VQPTLFAVMISLAALMRSGVPCAVLGHSLGEIAAAHYSGSLSDAARVYTLMSQAQT	122
Db	614	VQPALFAVMISLAALMRSHGVEPALVVGHSQGEIAAAHVAGALTLEDAAKLVAVRSRLR	673
Qy	123	TLAGTGLVSAVATDELLPRLPMTEDNPARLAVAANVGPS--TVSGAREVADLVAD	181
Db	674	RLGGGGGNASFLGTEQOAAERIGRPA---GALSIASVNGPSPVVVWAGESGLDELIAE	729
Qy	182	LTAAGVTRMTIPVDVPAHSPLMYAIEERVVSGLPITPRPSRIPFHSVYTGRLDTRELD	241
Db	730	CEAEKHKARRIPVDVASHSPQVESLRBELLTLAGISPSADVALYSTTTGQPIDATMD	789
Qy	242	AAYWYRNNSITVFEPPAARLLAQOGPKTEVMSPHPLVTLMGLQELAPLDGDTGTAD--T	299
Db	790	TAYWYANLREQVRFQDATRQLAEGPDAVVEVSPHVLTVGIEATL---DSALPADAGA	845
Qy	300	VIMGTLRGGCTLDHP	315
Db	846	CVGTLRRDRGGADF	861

RESULT 5

MCAS_MYCBO	STANDARD;	PRT;	2110 AA.
AC	Q02251;		
DT	01-JUN-1994 (Rel. 29, Created)		
DT	01-JUN-1994 (Rel. 29, Last sequence update)		
DT	15-JUL-1998 (Rel. 36, Last annotation update)		
DE	Mycocerosic acid synthase.		
GN	MAS.		
OS	Mycobacterium bovis.		
OC	Bacteria; Actinobacteria; Actinobacteriia (class); Actinobacteridae;		
OC	Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.		
OX	NCBI_TaxId=1765;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RX	MEDLINE=92406887; PubMed=1527058;		
RA	Maehur M., Kojatukudy P.E.;		
RT	"Molecular cloning and sequencing of the gene for mycocerosic acid		
RT	synthase, a novel fatty acid elongating multifunctional enzyme, from		
RT	Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guerin.",		
RL	J. Biol. Chem. 267:19388-19395(1992).		
CC	-1- FUNCTION: CATALYZES THE ELONGATION OF N-FATTY ACYL-COA WITH		
CC	METHYLMALONYL-COA (NOT MALONYL-COA) AS THE ELONGATING AGENT TO		
CC	FORM MYCOCEROSYL LIPIDS.		
CC	-1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTETHEINE.		
CC	-1- SUBUNIT: HOMODIMER WHOSE MONOMERS PROBABLY HAVE A HEAD TO TAIL		
CC	ARRANGEMENT.		
CC	-1- SUBCELLULAR LOCATION: Membrane-associated.		

CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, co-creation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/1db-sib.ch>).
CC or send an email to license@1db-sib.ch.

[illegible][illegible]

OS Mycobacterium tuberculosis.
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 CC NCBI_TaxID=1773;
 RX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=H37Rv;
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gae S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Kirogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey B.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains." to the EMBL/GenBank/DBJ databases.
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POLYPEPTIDE SYNTHASE INVOLVED IN PHENOLPTHIOCEROL
 CC SYNTHESIS.
 CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETHEINE (By
 CC similarity).
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL021070; CAA15929.1;
 DR EMBL; AE007122; AAK47329.1;
 DR TIGR; MT3002;
 DR TUBERCULAST; RV2932;
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantane attach.
 DR Pfam; PF00109; ketoacyl-synt. 1.
 DR Pfam; PF00550; pp-binding. 1.
 DR Pfam; PF00698; acyl_transf. 1.
 DR Pfam; PF02801; ketoacyl-synt. C. 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; FALSE_NEG.
 DR PROSITE; PS50075; ACP DOMAIN; 1.
 DR Hypothetical protein; Multifunctional enzyme; Oxidoreductase;
 DR Transferrase; NADP; Phosphopantetheine; Complete proteome.
 FT DOMAIN 1425 1495 ACRYL CARRIER (ACP).
 FT BINDING 1458 1458 PHOSPHOPANTHETHEINE (POTENTIAL).
 SO SEQUENCE 1538 AA; 162527 MW; B55E2A042AD00CC CRC64;
 Query Match 25.3%; Score 410; DB 1; Length 1538;
 Best Local Similarity 34.1%; Pred. No. 3.8e-22;
 Matches 109; Conservative 56; Mismatches 125; Indels 30; Gaps 10;
 Oy 1 VFVPGGQPGWPGGRELIDASDVFSVRACEAAPVYDVMSVEQVLRDSDPAPGIDRV 60
 Db 561 VFVPGGQPGWPGGRELIDASDVFSVRACEAAPVYDVMSVEQVLRDSDPAPGIDRV 619
 Oy 61 DVVPTLFVAVMISLAALMRSGVPECAVLGHSLGIEAAHVSGLSLADARVVTLSQA 120

Db 620 --10GLIGIQLTILTELMRSYGVOPDLVIGHSMGEVAAVAGALTFPABGLRVATPARL 677
 Oy 121 QTLTAGTG--ALYSVATPDELLPRIAPWTEEDNPARLAVAAVNGPSTVSGAREANADL 178
 Db 678 MAPLSGGGSMALLGLDAAATEAL-----IADVP-QVTGIVNSPRTVLAGPTEQIDEL 730
 Oy 179 VADLTAAQVTRMIPVDPVPHSPPLMVAIEERVVSGLLPITPRPSRTPFHSSTVGGRLDTR 238
 Db 731 IAVRAQNRASRVNIEVAHPNPMADLQPMRSELADLTFRPTTIGI-ISTTYADLHTQ 789
 Oy 239 EL--DAAYVNMSTVREFPARLLIQGP-----KTEVMSPPHYLTWGLQELADPLG 291
 Db 790 PIFDAEHWATNMENPVFQQA---IASGSGADGAYHTFIEISHPLLT---QAIDTLE 843
 Oy 292 D-----TTGTADTYIMGTLR 307
 Db 844 DAHPTKSAKRYLSIGTLQR 863
 RESULT 7
 PPSA_MYCTU STANDARD; PRT; 1876 AA.
 ID PPSA_MYCTU
 AC Q10977;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenolpthiocerol synthesis polyketide synthase ppsa.
 GN PPSA OR RV2931 OR MT3000 OR MTCY338.20.
 OS Mycobacterium tuberculosis.
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 CC NCBI_TaxID=1773;
 RX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=H37Rv;
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gae S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Kirogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey B.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains." to the EMBL/GenBank/DBJ databases.
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POLYPEPTIDE SYNTHASE INVOLVED IN PHENOLPTHIOCEROL
 CC SYNTHESIS.
 CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETHEINES (By
 CC similarity).
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Z74697; CAA98988.1; -.
 DR EMBL; AE007122; AAK47328.1; -.
 DR TIGR; MT3000; -.
 DR Tuberculin; RV2931; -.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantine_attach.
 DR Pfam; PF00109; ketoacyl-synt. 1.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl transf. 1.
 DR Pfam; PF02801; ketoacyl-synt C; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
 DR PROSITE; PS00606; B_KETOACYL SYNTHASE; 1.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 KM Multifunctional enzyme; Oxidoreductase; Transferase; NADP; Repeat;
 KM Phosphopantetheine; Complete proteome.
 FT DOMAIN 7 80
 FT DOMAIN 1764 1833
 FT NP_BIND 1498 1503
 FT BINDING 62 62
 FT ACT_SITE 273 273
 FT ACT_SITE 720 720
 FT BINDING 1796 1796
 FT BINDING 624 624
 FT CONFLICT 877 877
 FT CONFLICT 1323 1323
 SQ SEQUENCE 1876 AA; 198834 MW; D9783DBD48792110 CRC64;

Query Match 24.6%; Score 398; DB 1; Length 1876;
 Best Local Similarity 33.6%; Pred. No. 3.6e-21;
 Matches 107; Conservative 53; Mismatches 128; Indels 30; Gaps 9;

QY 1 VFVFGGQPMGPGMRELLDASDVFRSVCACAPAYVDWSVEQVLRDSDPAAGLDY 60
 632 VFVFGGSGQMGAGRGRLADPRFAAVALLEVFYEQAGFSLHDVLANEELVGLIQ- 690
 QY 61 DVVQPTLFAMWISLAALMRSQGVPCAVLGHSIGELIAAHVSGSLADAAVVTLSQA 120
 691 --IGLGIGMQLATLMLCSYGVNPDVLIGHSGEVAVAAGALTPEGLRATVATSRLL 748
 QY 121 OTTLAGTALVSVAATDELLPRIAPTE---DNPARLAVAANGPRSTVSGAREVA 176
 749 MAPSGGGGNA-----LLELDAPTEALADFP-QVTIGIYNSPROTVIAGPTQID 799
 QY 177 DLVADLTAAQVTRTMIPVDVPAHSPLMYAIEERVSGSLPTTPSPRIPHSSVTGRLD 236
 800 ELIRKVRQNRFPASRVVIEVAPHNPMADALQPARSELDLTPRTPTIGI-ISTTVADLH 858
 QY 237 TREL-DAVWYRNNSSTVRFEPAARLLQGP-----KTFVENSPPHVLTMGLQELAPD 289
 859 TQPVFDAEHMATNNRNPRFQQA---IASAGSGADGAVHTFELISAPHLTL---QAIIDT 912
 QY 290 LGDTTGTADTVIMGTLLR 307
 913 LHSAPGARAYTSLGTQR 930

RESULT 8
 MSAS_PENPA STANDARD; PRT; 1774 AA.
 AC P23367;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DS 6-methylsalicylic acid synthase (EC 2.3.1.-) (MSAS).
 OS Penicillium patulum (Penicillium griseofulvum).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; penicillium.
 NCBI_Taxid=5078;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=DSM 62862;
 RX MEDLINE=9106137; PubMed=2209605;

RA Beck J., Ripka S., Siegner A., Schlitz E., Schweizer E.;
 RT "The multifunctional 6-methylsalicylic acid synthase gene of
 RT Penicillium patulum. Its gene structure relative to that of other
 RT polyketide synthases";
 RL Eur. J. Biochem. 192:487-498(1990).
 CC -1- FUNCTION: This multifunctional enzyme is a polyketide synthase.
 CC It catalyzes a total of 11 steps by seven different component
 CC enzymes, in the biosynthesis of the antibiotic patulin.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + 3 malonyl-CoA + NADPH =
 CC 6-methylsalicylic acid + NADP(+) + 3 COA + 3 CO(2) + H(2)O.
 CC -1- PATHWAY: Patulin biosynthesis.
 CC -1- SUBUNIT: HOMOMULTIMER.
 CC -1- INDUCTION: IN THE LATE LOGARITHMIC GROWTH PHASE.
 CC -1- SIMILARITY: WITH MOST POLYKETIDE SYNTHASES, SPECIALLY WITH RAT
 CC FATTY ACID SYNTHASE, AND WITH OTHER ENZYMES SUCH AS LIPASES AND
 CC THIOLEASES.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; X55776; CAA39295.1; -.
 DR PIR; S13178; S13178.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantine_attach.
 DR Pfam; PF00109; ketoacyl-synt. 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR Pfam; PF00698; Acyl transf. 1.
 DR Pfam; PF02801; ketoacyl-synt C; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
 DR PROSITE; PS00606; B_KETOACYL SYNTHASE; 1.
 DR PROSITE; PS50075; ACP DOMAIN; 1.
 KM Multifunctional enzyme; Oxidoreductase; Antibiotic biosynthesis;
 KM Transferase; NADP; Phosphopantetheine.
 FT DOMAIN 186 238
 FT DOMAIN 642 676
 FT NP_BIND 1403 1450
 FT DOMAIN 1700 1769
 FT NP_BIND 1419 1424
 FT ACT_SITE 204 204
 FT ACT_SITE 653 653
 FT BINDING 1732 1732
 SQ SEQUENCE 1774 AA; 190732 MW; 05BD5DD10863F938 CRC64;

Query Match 22.7%; Score 366.5; DB 1; Length 1774;
 Best Local Similarity 31.0%; Pred. No. 6.6e-19;
 Matches 100; Conservative 64; Mismatches 132; Indels 27; Gaps 8;

QY 1 VFVFGGQPMGPGMRELLDASDVFRSVCACAPAYVDWSVEQVLRDSDPAAGLDY 60
 567 VFVFGGAGQMGPMGKQLIH-NPVFPAALQPLDELIGALGSLTELLR-TGDPRESSDRV 624
 QY 61 DVVQPTLFAMWISLAALMRSQGVPCAVLGHSIGELIAAHVSGSLADAAVVTLSQA 120
 625 QIL---TYVMQIGLSALOSNGITPQAVIGHSGEIAASVAGALSFAEGALVITRRALL 681
 QY 121 OTTLAGTALVSVAATDELLPRIAPTEDNPARLAVAANGPRSTVSGAREVAD 177
 682 YRQVMGKGMIIVNLVLSAEETEL-----GSRSLVVAIINSSPSCVAGKELVAE 733
 QY 178 LVADLTAAQVTRTMIPVDVPAHSPLMYAIEERVSGSLPTTPSPRIPHSSVTG 233
 734 TAELTKARVGTFTVSDIAFHSPTLNGVLPRLDVLAEFTLSPVSP---NVKLYSTALAD 790
 QY 234 RLDTRELDAAVWYRNNSSTVRFEPAARLLQGPKTFVENSPPHVLTMGLQELAPD 293
 791 PRGQDLADVEYMGANNVNRVLTSAVAAVEDGYRLFLVSTHPPVSHSINETLMD---- 846

DB	QY	294	TGTA	DTVM	MG	TLR	GC	GT	HL	DF	HL	316
DB			847	AGMED	PAVIT	PTL	LRK	PK	PT	ER	GILL	869

RESULT 9

ID	FABD	BACSU	STANDARD;	PRT,	317	AA.
AC	P71019;	Q34463;				
DT	01-NOV-1997	(Rel. 35,	Created)			
DT	30-MAY-2000	(Rel. 39,	Last sequence update)			
DT	15-JUN-2002	(Rel. 41,	Last annotation update)			
DE	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39) (MCT).					
GN	FABD.					
OS	Bacillus subtilis.					
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillue.					
OX	NCBI_TaxID=1423;					
RN	SEQUENCE FROM N.A.					
RC	STRAIN=168;					
RX	MEDLINE=96326321; PubMed=8759840;					
RA	Morbidioti H.R., de Mendoza D., Cronan J.E. Jr.;					
RT	"Bacillus subtilis acyl carrier protein is encoded in a cluster of					
RL	lipid biosynthesis genes."					
RN	J. Bacteriol. 178:4794-4800 (1996).					
RP	(2)					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=168;					
RX	MEDLINE=98195738; PubMed=9534248;					
RA	Foulinger D., Errington J.;					
RT	"A 28 kbp segment of the spvM region of the Bacillus subtilis 168					
RL	genome."					
RL	Microbiology 144:801-805 (1998).					
RN	(3)					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=168;					
RX	MEDLINE=98044033; PubMed=9384377;					
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,					
RA	Azevedo V., Bertero M.G., Besiieres P., Bolotin A., Borchart S.,					
RA	Borries R., Boutsier L., Brans A., Braun M., Brignell S.C., Bron S.,					
RA	Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,					
RA	Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,					
RA	Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,					
RA	Enrian K.D., Errington J., Fabre C., Ferrari E., Founger D.,					
RA	Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,					
RA	Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,					
RA	Guiseppi G., Guy B.J., Haea K., Halech J., Harwood C.R., Henaut A.,					
RA	Hilbert H., Holshappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,					
RA	Joris B., Karamata D., Kasahara Y., Klaber-Blanchard M., Klein C.,					
RA	Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,					
RA	Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,					
RA	Lee S.W., Levine A., Liu H., Maesuda S., Manuel C., Medigue C.,					
RA	Medina N., Meliada R.P., Mizuno M., Moesti D., Nakai S., Noback M.,					
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,					
RA	Parro V., Pohl T.M., Portetelle D., Portwollik S., Prescott A.M.,					
RA	Presanen E., Pujic P., Punnelle B., Rapoport G., Rey M., Reynolds S.,					
RA	Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,					
RA	Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,					
RA	Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,					
RA	Tacchini A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,					
RA	Tateuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,					
RA	Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassaretti A.,					
RA	Viari A., Wambut R., Wedler E., Wedler H., Weitzner T.,					
RA	Winters P., Wipet A., Yamamoto H., Yamane K., Yasunoco K., Yata K.,					
RA	Yoshida K., Yoshikawa H.F., Zumestein E., Yoshikawa H., Danchin A.;					
RT	"The complete genome sequence of the Gram-positive bacterium Bacillus					
RL	subtilis."					
RL	Nature 390:249-256 (1997).					
CC	-1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +					
CC	malonyl-[acyl-carrier protein].					
CC	-1- PATHWAY: Fatty acid biosynthesis.					
CC	-1- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH					
CC	TRANSACYLASE ACTIVITY.					

-- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/termsandconditions or send an email to license@ebi.ac.uk).

DR EMBL; U59833; AAC04306.1; --
DR EMBL; Y13537; CAAT4249.1; --
DR EMBL; Z99112; CAB13463.1; --
DR HSPB; P25715; IMMA.
DR Subtilisin; BG1836; fabd.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR004410; FAbD
Plam; PF00696; ACPyl.transf.; 1.
DR TIGRPFMS; BIOR00128; fabd; 1.
KW Fatty acid biosynthesis; Transferase; Complete proteome.
FT ACT_SITE 91 91 BY SIMILARITY.
FT SITE 201 201 BY SIMILARITY.
FT CONFLICT 124 131 VPAGEGM -> GGRLAKEW (IN REF. 1).
SQ SEQUENCE 317 AA; 3403 MW; 191A828B81C9IDTF CXC64;

Query Match 16.4%; Score 265; DB 1; Length 317;
Best Local Similarity 28.3%; Pred. No. 2.1e-12;
Matches 83; Conservative % 50; Mismatches 130; Indels 30; Gaps 8;

DY 2 FVPEGQGPMPNGMKRELDD----ASDVRESVRACEAAFAFYDMVSQGVLRLSDPADGL 57
E FLFPGGGSGPTFGMCKELXYBPAPKRPFLDEADLETLETKLSLI-----PBGDAHEL 56

DY 58 DRVAVVPCTPAFWISALIAWRSSGVPCAVLGHSIGEIAAHYSGGSLADAARVLTLM 117
S TLTVNAQPALLTTTSIVALEPKESGIITPDTPAGHSLGEYSALTVAAGALSFPDAYTYVKRR 116

DY 118 SQ--AQOTTLTAGTALSVVANTPPDELFRIPAWTEEDNPARRALAANAANGSRSTVGAREAY 175
I GFENNEAVEPVPAEGEMAAAILGMDADALKOVTDKYTEENGNLVOLANINCPOGQVIISTAGNAV 176

DY 176 ADLVADVATAAOVRTPMIPVDV--PAHSPLMWAIIEERVVSGLLPIPTRPSRIFFSHSYTG 233
S ELASELKAKENGAKRALPLEVSGFPHSELMKPRAEKIKEVLADACDIKDADVPVISNVSAD 235

DY 234 ---RLDTRE-LDAAYWRNMSTVFEPPEARILLQQGPKTEVMSPHPVLT 280
S VMTEKADIKKETILEOLY-----SPVAFEEISINKLIAEGVTTFIIEGKVLS 282

RESULT 10

ID WA_EMENT STANDARD; PRU; 1986 RA.

WA_EMENT AC Q03149;
DT 01-JUN-1994 (Rel. 29, Created)
RT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Conidial green pigment synthase (EC 2.3.1.-).
GN WA.
OS Emerticella nidulans (Aspergillus nidulans).
OC Eurotiaceae; Phungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
CC Eurotiales; Trichocomaceae; Emerticella.
OX NCBI_TaxId=5072;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93101122; PubMed=1465094;
RA Mayorga M.E., Timberlake W.E.;
RT "The developmentally regulated Aspergillus nidulans wa gene encodes a polyketide homologous to polyketide and fatty acid synthases";,
ML Mol. Gen. Genet. 235:205-212(1992)
FN FUNCTION: THIS PROTEIN CONDENSES CARBON UNITS TO FORM AN INTERMEDIATE YELLOW POLYKETIDE PIGMENT THAT IS POLYMERIZED BY CONIDIAL LACCASE TO FORM THE GREEN PIGMENT IN MATURE ASEXTUAL SPORES (CONIDIA).

CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES
(POTENTIAL).
CC -1- PATHWAY: Conidial green pigment biosynthesis.
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ldb-sib.ch/announce/>
CC or send an email to license@ldb-sib.ch).

CC EMBL; X65866; CAA6695.1; -
DR PIR; S28353; S28353.
DR Interpro; IPR001227; Ac transferase.
DR Interpro; IPR000794; Ketoacyl-synt.
DR Interpro; IPR003880; Ppanne_attach.
DR Interpro; IPR001031; Thioesterase.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF00550; pp-binding; 2.
DR Pfam; PF00698; Acyl-transf; 1.
DR Pfam; PF00975; Thioesterase; 1.
DR Pfam; PF02801; ketoacyl-synt_C; 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR PROSITE; PS00606; B KETOACYL SYNTHASE; 1.
DR PROSITE; PS50075; ACP DOMAIN; 2.
KM Transferase; Phosphopantetheine; Multifunctional enzyme; Repeat.
FT DOMAIN 529 582 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT DOMAIN 991 1024 ACYL/MALONYL TRANSFERASES
(BY SIMILARITY).
FT DOMAIN 1650 1719 ACYL CARRIER (ACP) 1.
FT DOMAIN 1772 1841 ACYL CARRIER (ACP) 2.
FT ACT_SITE 548 548 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT ACT_SITE 1001 1001 ACYL/MALONYL TRANSFERASES (BY
SIMILARITY).
FT BINDING 1682 1682 PHOSPHOPANTETHEINE (BY SIMILARITY).
FT BINDING 1804 1804 PHOSPHOPANTETHEINE (BY SIMILARITY).
SQ SEQUENCE 1986 AA; 216634 MW; 74EF0940FP40EB9A CRC64;
Query Match 15.9%; Score 258; DB 1; Length 1986;
Best Local Similarity 27.2%; Pred. No. 5.8e-11;
Matches 90; Conservative 54; Mismatches 139; Indels 48; Gaps 10;

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39) (MCT).
GN FMBD OR STM1194.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxId=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L72;
RC MEDLINE=98317265; PubMed=9642179;
RA Zhang Y., Cronan J.E. Jr.;
RT "transcriptional analysis of essential genes of the Escherichia coli
RT fatty acid biosynthesis gene cluster by functional replacement with
RT the analogous Salmonella typhimurium gene cluster".
RL J. Bacteriol. 180:3295-3303(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=L72 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "complete genome sequence of Salmonella enterica serovar Typhimurium
RT L72".
RL Nature 413:852-856(2001).
CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
malonyl-[acyl-carrier protein].
CC -1- PATHWAY: Fatty acid biosynthesis.
CC -1- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH
TRANSACYLASE ACTIVITY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ldb-sib.ch/announce/>
CC or send an email to license@ldb-sib.ch).

CC EMBL; AF04668; AAC38649.1; -
DR EMBL; AE008752; AL20123.1; -
DR HSP; F25715; IMLA.
DR ScyGene; SG10637; fadb.
DR Interpro; IPR001227; Ac transferase.
DR Interpro; IPR004410; fadb.
DR Pfam; PF00698; Acyl-transf; 1.
DR TIGRfam; TIGR00128; fadb; 1.
KW Fatty acid biosynthesis; transferase; Complete proteome.
FT INIT_MET 0 BY SIMILARITY.
FT ACT_SITE 91 91 BY SIMILARITY.
FT ACT_SITE 200 200 BY SIMILARITY.
SQ SEQUENCE 308 AA; 32274 MW; B3565B45341A611A CRC64;
Query Match 15.8%; Score 255; DB 1; Length 308;
Best Local Similarity 28.6%; Pred. No. 1.e-11;
Matches 90; Conservative 50; Mismatches 141; Indels 34; Gaps 11;

QY 177 DLVADLTAAGVTEM-IPVDPAHSPMLVAIEERVSGLLPTPRSPRIPPHSSVTGRL 235
 DB 177 RAGAACCAAGAAKRLPLPVSPSHCALMKPRADLALVELAKITTSATFVPMNV----- 231
 QY 236 DTR-ELDAA---WYRNMSTVFEEPARLLLOQGFRTVEMSPHPVLTWGLQELAPDL 290
 DB 232 DVKCEIDAAAIKDALVRLVQLVNPVQWTKSVFETIAQGVHEHVEVGPVKLT----- 281
 QY 291 GDTTGTADTVIMGTL 305
 DB 282 GLTKRIIVDTLTASAL 296

RESULT 12
 PABD_ECOLI
 ID PABD_ECOLI STANDARD; PRT; 308 AA.
 AC P25715;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39) (MCT).
 GN PABD OR TPA OR B1092.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN 11
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
 RC STRAIN=K12;
 RX MEDLINE=92183950; PubMed=1339356;
 RA Magnuson K., Oh W., Larson T.J., Cronan J.E. Jr.;
 RT "Cloning and nucleotide sequence of the fabD gene encoding malonyl
 coenzyme A-acyl carrier protein transacylase of Escherichia coli.";
 RL FEBS Lett. 299:262-266(1992).
 RN 12
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92234941; PubMed=1314802;
 RA Verwoert I.J., Verbrée E.C., van der Linden K.H., Nijkamp H.J.;
 RT "Cloning, nucleotide sequence, and expression of the Escherichia coli
 fabD gene, encoding malonyl coenzyme A-acyl carrier protein
 transacylase.";
 RL J. Bacteriol. 174:2851-2857(1992).
 RN 13
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95214624; PubMed=7700236;
 RA Bouguin N., Tempete M., Holland I.B., Seror S.-J.;
 RT "Resistance to trifluoperazine, a calmodulin inhibitor, maps to the
 fabD locus in Escherichia coli.";
 RL Mol. Gen. Genet. 246:628-637(1995).
 RN 14
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RN 15
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajinara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Mochimaru K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horikuchi T.;
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 12.7-28.0 min region on the linkage map.";

RL DNA Res. 3:137-155(1996).
 RN 16
 RP SEQUENCE OF 288-308 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=92210530; PubMed=1556094;
 RA Cronan J.E. Jr., Rawlings M.;
 RT "The gene encoding Escherichia coli acyl carrier protein lies within
 a cluster of fatty acid biosynthetic genes.";
 RL J. Biol. Chem. 267:5751-5754(1992).
 RN 17
 RP SEQUENCE OF 1-11.
 RC STRAIN=K12 / W3110;
 RA Pasquali C., Sanchez J.-C., Ravier F., Golaz O., Hughes G.J.,
 RA Frutiger S., Paquet N., Wilkins M., Appel R.D., Bairoch A.,
 RA Hochstrasser D.F.;
 RL Submitted (SEP-1994) to the SWISS-PROT data bank.
 RN 18
 RP SEQUENCE OF 1-10.
 RC STRAIN=K12 / EMG2;
 RX MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robison K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 in the genome of Escherichia coli K-12";
 RL Electrophoresis 18:11259-1313(1997).
 RN 19
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
 RX MEDLINE=95286570; PubMed=7768883;
 RA Serre L., Verbrée E.C., Dauter Z., Stuitje A.R., Derewenda Z.S.;
 RT "The Escherichia coli malonyl-CoA:acyl carrier protein transacylase
 at 1.5-A resolution. Crystal structure of a fatty acid synthase
 component.";
 RL J. Biol. Chem. 270:12961-12964(1995).
 CC 1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 malonyl-[acyl-carrier protein].
 CC 1- PATHWAY: Fatty acid biosynthesis.
 CC 1- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH
 TRANSACYLASE ACTIVITY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M87040; AAA23742.1; -;
 DR EMBL; Z11565; CA477658.1; -;
 DR EMBL; M84991; AAA23738.1; -;
 DR EMBL; AE000210; AAC74176.1; -;
 DR EMBL; D90745; BAA35900.1; -;
 DR PIR; S20443; S20443.
 DR PIR; B41856; B41856.
 DR PDB; 1MLA; 25-JAN-95.
 DR SWISS-2DPAGE; P25715; COI1.
 DR EcoGene; EG11317; fabD.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR004410; FabD.
 DR Pfam; PF00698; Acyl_transf_1.
 DR TIGRFAMs; TIGR00128; fabD; 1.
 KW Fatty acid biosynthesis; Transferase; 3D-structure; Complete proteome.
 FT INIT MET 0
 FT ACT_SITE 91 91
 FT ACT_SITE 200 200
 FT ACT_SITE 308 AA, 3286 MM, F5901043D92FED8E CRC64;
 SQ SEQUENCE

Query Match 15.6%; Score 252; DB 1; Length 308;
 Best Local Similarity 27.9%; Pred. No. 1.8e-11;
 Matches 88; Conservative 49; Mismatches 144; Indels 34; Gaps 10;
 2 FVFPQGPQPMGKRELLSDVFRSVPACAEAPFYDWSVEQVLRDPSDPAFGDVRD 61
 DB 5 FVFPQGSQTVGKMLADMAASYPVEETFAEASALG-YDLMALTO-----QGPAEINKTW 59


```

CC -1- PATHWAY: Sterigmatocystin biosynthesis; first step.
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its use by
CC the European Bioinformatics Institute. There are no restrictions on its use by
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce
CC or send an email to license@isb-sdb.ch).
CC
CC EMBL; U34740; AAC49191.1; -.
CC EMBL; L39121; AAA81586.1; -.
CC HSSP; P25715; IMEA
CC
CC InterPro; IPR001227; Ac_transferase.
CC InterPro; IPR000794; Ketoacyl-synt.
CC InterPro; IPR003880; Pantine atach.
CC InterPro; IPR001031; Thioesterase.
CC Pfam; PF00109; ketoacyl-synt; 1.
CC Pfam; PF00550; pp-binding; 2.
CC Pfam; PF00698; Acyl_transf; 1.
CC Pfam; PF00975; Thioesterase; 1.
CC Pfam; PF02801; ketoacyl-synt C; 1.
CC PROSITE; PS00075; ACP DOMAIN; 2.
CC Transferrase; Acyltransferase; Phosphopantetheine; Repeat;
CC Multifunctional enzyme.
CC
CC FT DOMAIN 383 814 BETA-KETOACYL SYNTHASE.
CC FT DOMAIN 864 1209 ACYL/MALONYL TRANSFERASES.
CC FT DOMAIN 1706 1777 ACYL CARRIER (ACP) 1.
CC FT DOMAIN 1830 1901 ACYL CARRIER (ACP) 2.
CC FT DOMAIN ? 2181 THIOESTERASE.
CC FT ACT_SITE 552 552 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
CC FT ACT_SITE 978 978 ACYL/MALONYL TRANSFERASES (BY SIMILARITY).
CC
CC FT BINDING 1738 1738 PHOSPHOPANTHETHEINE (BY SIMILARITY).
CC FT BINDING 1862 1862 PHOSPHOPANTHETHEINE (BY SIMILARITY).
CC FT ACT_SITE 2028 2028 THIOESTERASE (BY SIMILARITY).
CC
CC SQ SEQUENCE 2181 AA; 238831 MW; 5A3B5712AA9AD942 CRC64;
CC
CC Query Match 15.0%; Score 243.5; DB 1; Length 2181;
CC Best Local Similarity 27.8%; Pred. No. 7.3e-10;
CC Matches 90; Conservative 56; Mismatches 153; Indels 25; Gaps 8;
CC
CC QY 1 VFVPRGCGPQMPFGNGRELLDSDVFRSEVRCGEAAFAFYDWSVEQVLRDSPDAFGLDLV 60
CC Db VIACSGGSGSQTGCGMWOLYNSYPTFRSDLERFDOLARSYGPPSELYVTSKPVGDSMEDL 944
CC
CC QY 61 --DVVQPLTFVMTSIALWMSQGVPCAVYGHSGEYIAAAHVSQGLSLAARVYTLWS 118
CC Db LPVIVQLALVLEHMLGNLWSFGKRPAAVYGHSGEYIAALYISGVLSAAOTLYLVGKRA 1004
CC
CC QY 119 Q--AQTTLAAGALVAVNATPDELLPRIAPTEDENPALIAVAANVGPRSTVVGAREAVA 176
CC Db KILGERCGRGTHAMLAVASPVTLCEVLA-----ESN-----CEVAHNGNPDTVLSGLKQVM 1058
CC
CC QY 177 DLVADLTAAQVRTKMIPIVDVAHSELMVAIEERVVSGLLPTTPRSRIPEHSSVTGGRLD 236
CC Db NLQNSLSATYGIKGLTLKLPFAHSAQVQPIEEFRNVAAGVTFPKPQIPIVLSPLLVKVID 1118
CC
CC QY 237 TR-ELDAAYWRNMSSTVRFEPFA--AR-----LLDQGPRTVFMSPHYLTIMG 283
CC Db EKGTVADPYTLARHCRPEKPVKMSVLEHADDOHIIITDRTIVIVGPKALMAGMIKTTLDKOT 1178
CC
CC QY 284 QELAPDLDGTTGT--ADTVINGTL 305
CC Db 1179 SSALPTLGPISLDWKSLSLTNLTGL 1202
CC
CC RESULT 15
CC FAAD SYNY3 STANDARD; PRT; 293 AA.
CC ID FAAD SYNY3
CC NC P73242;
CC DT, 01-NOV-1997 (Rel. 35, Created)

```

[illegible]

Search completed: June 17, 2003, 13:02:31
Job time : 5:44608 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 17, 2003, 12:55:32 ; Search time 7.05626 Seconds
(without alignment)
4305.183 Million cell updates/sec

Title: US-09-914-286-3_COPY_29_344

Perfect score: 1618
Sequence: 1 VFVPPGQGPQMPGKRELLD.....ADTVMGTLRRGGTLDHFL 316

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match: 0%
Maximum Match: 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	817.5	50.5	2124	2 T28658	polyketide synthase
2	773	47.8	4151	2 G70944	probable polyketid
3	771.5	47.7	3573	2 S23070	erythronolide synt
4	722	44.6	3491	2 T43231	probable 6-deoxyer
5	707	43.7	3519	2 S43048	polyketide synthase
6	699	43.2	1937	2 T03224	probable polyketid
7	697.5	43.1	5069	2 T17464	rifamycin polyketi
8	696.5	43.0	4613	2 T17409	polyketide synthase
9	689.5	42.6	2126	2 H70621	probable polyketid
10	689	42.6	3172	2 S22012	erythronolide synt
11	689	42.6	3178	2 S13595	6-deoxyerythronol
12	681.5	42.1	3739	2 T17410	polyketide synthase
13	674	41.7	1562	2 T17411	polyketide synthase
14	673.5	41.6	1763	2 T17465	rifamycin polyketi
15	667.5	41.3	1346	2 T17412	polyketide synthase
16	661	40.9	1728	2 T17466	rifamycin polyketi
17	658	40.7	4735	2 T17463	rifamycin polyketi
18	649.5	40.1	3413	2 T17467	rifamycin polyketi
19	638	39.4	7576	2 T17428	PKS56 polyketide s
20	600	37.1	2723	2 T03221	probable polyketid
21	595	36.8	2100	2 T03223	probable polyketid
22	578	35.7	8563	2 T30226	polyketide synthase
23	566	35.0	6260	2 T30228	polyketide synthase
24	553	34.2	10223	2 T30225	polyketide synthase
25	539	33.3	6420	2 T30283	polyketide synthase
26	516.5	31.9	1570	2 AC2012	hypothetical prote
27	502.5	31.1	1582	2 E70876	probable polyketid
28	502	31.0	2108	2 H70819	probable polyketid
29	479	29.6	2126	2 E70522	probable polyketid

30	467.5	28.9	2111	2 A70668	mycocerosate synth
31	466.5	28.8	2118	2 S72705	mycocerosate synth
32	463.5	28.6	2110	2 B44110	mycocerosate synth
33	460.5	28.5	2116	2 C86926	probable mycoceros
34	457.5	28.3	1293	2 T30871	orrellinic acid sy
35	441.5	27.3	1827	2 B70984	probable polyketid
36	440	27.2	2458	2 T17420	hypothetical prote
37	437	27.0	1587	2 AB2012	hypothetical prote
38	436	26.9	2297	2 T34918	polyketide synthase
39	432.5	26.7	1795	2 D97312	hypothetical prote
40	429	26.5	1395	2 S18954	fix23-2 protein -
41	427	26.4	1017	2 B70985	probable polyketid
42	417	25.8	635	2 S73017	polyketide synthase
43	417	25.8	1446	2 S73013	polyketide synthase
44	417	25.8	1540	2 H87203	polyketide synthase
45	417	25.8	1871	2 A87204	polyketide synthase

ALIGNMENTS

RESULT 1

T28658
polyketide synthase - Sorangium cellulosum (fragment)
C/Species: Sorangium cellulosum
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 01-Dec-2000
C/Accession: T28658
R/Schupp, T.; Toupet, C.; Cluzel, B.; Neff, S.; Hill, S.; Beck, J.J.; Ligon, J.M.
J. Bacteriol. 177, 3673-3679, 1995
A/Title: A Sorangium cellulosum (myxobacterium) gene cluster for the biosynthesis of th
inomyces.
A/Reference number: Z20499; MUID:95325306; PMID:7601830
A/Accession: T28658
A/Status: preliminary; translated from GB/EMBL/DBD
A/Molecule type: DNA
A/Residues: 1-2124 <SCH>
A/Cross-references: EMBL:U24241; NID:9868150; PID:9868151; PIDN:AAA79984.1
A/Experimental source: strain So ce26
C/Function:
A/Description: involved in soraphen A biosynthesis
C/Superfamily: acyl carrier protein homology
C/Keywords: carrier protein; phosphopantetheine; phosphoprotein
F/335-406/Domain: acyl carrier protein homology <ACP>
F/370/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match	Score	817.5;	DB 2;	Length	2124;
Best Local Similarity	53.5%			Pred. No. 3.5e-53;	
Matches	169;			Mismatches	96;
				Indels	9;
				Gaps	2;
QY	1	VFVPPGQGPQMPGKRELLDASDVFPRESVACAAFPAYDMSVEOYLKDSPPAAGLDV	60		
DB	969	VFVPPGQGPQMPGKRELLDASDVFPRESVACAAFPAYDMSVEOYLKDSPPAAGLDV	1028		
QY	61	DVVOPTLFAVWISLAALMRSQVPCAVLGHSLGEIAAAHVSGLSLADAAVVTWLSQA	120		
DB	1029	DVVOPTLFAVWISLAALMRSQVPCAVLGHSLGEIAAAHVSGLSLADAAVVTWLSQA	1088		
QY	121	QTLTAGALVSAATPDELLPRIAPRTENPRALVAANGRSTVSGARAVADVVA	180		
DB	1089	LTTVAGGAAVAALVGLASDITLAPMGD---RLSTAAVNSPRATLVSGEPALIDLID	1144		
QY	181	DLTAQVTRMIEVDVPAHSPLMYAIEBRVSGLLPTPRSRIPFHSSTYGGDLTREL	240		
DB	1145	SLTPAQQFAKAVDVDAASHQAQNDVDELAAAGLANIAPRTCELPLVSTYTGRLDSSSL	1204		
QY	241	DAAYWRNNSSTVRFEPARLLLOQGEKTFVENSPPHVLTMGLQELAPDLGDTTGTADTV	300		
DB	1205	DGAYWYRNLRQTVLFSSATRLDDGHRFVEVSPHVLTLALRETC-----ERSPLDPV	1259		
QY	301	IMGTLRGGTLDHFL 316			
DB	1260	VVGSIIRDDEGLARLL 1275			

RESULT 2

709944
probable polyketide synthase Rv2048c - Mycobacterium tuberculosis (strain H37Rv)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C/Accession: G70944
R/Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Gentles, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: G70944
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-4151 <COL>
A/Cross-references: GB:AL021899; GB:AL123456; NID:93242282; PID:CAA17262.1; PID:9289678
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: pks12
C/Superfamily: Mycobacterium tuberculosis probable polyketide synthase Rv2048c; 3-oxoacyl homology; [acyl-carrier-protein] S-malonyltransferase homology
C/Keywords: carrier protein
F/56-455/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F/559-841/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
F/1680-1861/Domain: short-chain alcohol dehydrogenase homology <SAD1>
F/1963-2038/Domain: acyl carrier protein homology <ACP1>
F/2078-2478/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F/2582-2860/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F/3710-3891/Domain: short-chain alcohol dehydrogenase homology <SAD2>
F/3995-4066/Domain: acyl carrier protein homology <ACP2>

Query Match 47.8%; Score 773; DB 2; Length 4151;
Best Local Similarity 51.6%; Pred. No. 1.8e-49;
Matches 163; Conservative 45; Mismatches 102; Indels 6; Gaps 2;

1 VFVPPGGQPMGKRELLDASDVRESVACCAAFAPYVMSVQVLRDSDPAPGLDRV 60
559 VFVPPGGQPMGKRELLDASDVRESVACCAAFAPYVMSVQVLRDSDPAPGLDRV 618
61 DVVQPTLFAVMISLAALMRSGVPCAVLGHSIGETIAAAHVSGLSLDAAARVTLMSQA 120
619 DVVQPTLFAVMISLAALMRSGVPCAVLGHSIGETIAAAHVSGLSLDAAARVTLMSQA 678
121 QTTIAGTALVSVAATPELLPRIAPWTEBNPARLAANAANGPSTVSGAREVADIVA 180
679 LAGLAGPGMWSIACGADQARDLAPGCD---RVSIADVNGPSAVVSVGEVGLAEILIA 734
181 DLTAQVTRMIPVDVPAHSPMLMTAIBERVVSGLLPPTPPSRIPPHSSVTGRLDREL 240
735 VCSIKELTRRIEDVYASHVEKIRGPLEALSGIEPRSTRIVFSTVGNRLDTAGL 794
241 DAAVYRNMSSTVFEPPAARLLLOQGPKEVMSPPHVLTMGLQELAPDLGDTTGTADTV 300
795 DADVYRNVRQTVLPDQVNAACEGCVRTFESSPHRLITGVETTF--AACTGDSEAI 852
QY 301 IMGLRGGQTLDFHL 316
DB 853 VVPTLRGGDGLHRL 868

RESULT 3

S23070
erythronolide synthase (EC 2.3.1.94) II - Saccharopolyspora erythraea
N/Alternate names: 6-deoxyerythronolide B synthase II
C/Species: Saccharopolyspora erythraea
C/Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 21-Jul-2000
C/Accession: S23070; S22011; S23205
R/Bevitt, D.J.; Cortes, J.; Haydock, S.F.; Leadlay, P.F.
Eur. J. Biochem. 204, 39-49, 1992
A/Title: 6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of

A/Reference number: S23070; MUID:92155230; PMID:1740151

A/Accession: S23070

A/Molecule type: DNA

A/Residues: 1-3573 <BEV1>

A/Cross-references: EMBL:X62569

A/Experimental source: strain NRRL 2338

R/Bevitt, D.J.

submitted to the EMBL Data Library, September 1991

A/Reference number: S22011

A/Accession: S22011

A/Molecule type: DNA

A/Residues: 1-184,'T',186-301,'S',303-521,523-658,'A',660-993,1001-1212,'H',1214-1392,13479,'DH',3480-3572 <BEV2>

A/Cross-references: EMBL:X62569; NID:946977; PID:CAA44448.1; PID:9581651

R/Caffrey, P.; Bevitt, D.J.; Staunton, J.; Leadlay, P.F.

FEBS Lett. 304, 225-228, 1992

A/Title: Identification of DEBS 1, DEBS 2 and DEBS 3, the multienzyme polypeptides of t

A/Reference number: S23103; MUID:92316235; PMID:1618327

A/Accession: S23205

A/Molecule type: protein

A/Residues: 2-12,'XXX' <CAF>

A/Experimental source: strain CA340

C/Genetics:

A/Gene: eryA

A/Start codon: GTG

C/Function:

A/Description: catalyzes the construction of a polyketide chain, which is then cyclised

A/Superfamily: [acyl-carrier-protein] S-malonyltransferase homology; 3-oxoacyl-[acyl-ca

ort-chain alcohol dehydrogenase homology

C/Keywords: acyltransferase; antibiotic biosynthesis; carrier protein; coenzyme A; mult

F/52-453/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>

F/561-843/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>

F/1140-1308/Domain: short-chain alcohol dehydrogenase homology <SAD2>

F/1404-1475/Domain: acyl carrier protein homology <ACP1>

F/1519-1919/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>

F/2023-2305/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>

F/2857-3131/Domain: long-chain alcohol dehydrogenase homology <LADH>

F/3149-3327/Domain: short-chain alcohol dehydrogenase homology <SADH>

F/3420-3493/Domain: acyl carrier protein homology <ACP2>

F/1439/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 47.7%; Score 771.5; DB 2; Length 3573;
Best Local Similarity 49.7%; Pred. No. 2e-49;
Matches 157; Conservative 44; Mismatches 108; Indels 7; Gaps 2;

1 VFVPPGGQPMGKRELLDASDVRESVACCAAFAPYVMSVQVLRDSDPAPGLDRV 60
561 VFVPPGGQPMGKRELLDASDVRESVACCAAFAPYVMSVQVLRDSDPAPGLDRV 620
61 DVVQPTLFAVMISLAALMRSGVPCAVLGHSIGETIAAAHVSGLSLDAAARVTLMSQA 120
621 DVVQPTLFAVMISLAALMRSGVPCAVLGHSIGETIAAAHVSGLSLDAAARVTLMSQA 680
121 QTTIAGTALVSVAATPELLPRIAPWTEBNPARLAANAANGPSTVSGAREVADIVA 180
681 MRSLSGCGMAAVALGAARERLRPMQD---RLSVAANNGRSVSVSEPGALRAFSE 736
QY 181 DLTAQVTRMIPVDVPAHSPMLMTAIBERVVSGLLPPTPPSRIPPHSSVTGRLDREL 240
DB 737 DCAAGGIRVADIVDVASHSPQIERVAREELLETGDIAPAPAVTFPHSTVSRSMGTEL 796
241 DAAVYRNMSSTVFEPPAARLLLOQGPKEVMSPPHVLTMGLQELAPDLGDTTGTADTV 300
DB 797 DARIWYNLRERETVAFADAVTRLAESGYDAETEVSPHVVAQAVEA---VEEADGADAV 853
QY 301 IMGLRGGQTLDFHL 316
DB 854 VVGSIRHDDGLSFL 869

RESULT 4

T43231

probable 6-deoxyerythronolide B synthase, module 1/2 - Saccharopolyspora erythraea
C/Species: Saccharopolyspora erythraea
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 01-Dec-2000
C/Accession: T43231
R/Donado, S.; Stever, M.J.; McAlpine, J.B.; Swanson, S.J.; Katz, L.
Science 252, 675-679, 1991
A/Title: Modular organization of genes required for complex polyketide biosynthesis.
A/Reference number: 222354; MUID:91220065; PMID:2024119
A/Accession: T43231
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-3491 <DON>
A/Cross-references: EMBL:M3676; NID:9152691; PID:9152692; PIDN:AAA6493.1
C/Genetics:
A/Genes: eryA
C/Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protei
C/Keywords: carrier protein; phosphopantetheine; phosphoprotein
F/57-336/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F/416-483/Domain: acyl carrier protein homology <ACP1>
F/526-926/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F/1031-1317/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F/1886-1957/Domain: acyl carrier protein homology <ACP2>
F/2001-2400/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F/2508-2791/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
F/3332-3403/Domain: acyl carrier protein homology <ACP3>
F/1921/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 44.6%; Score 722; DB 2; Length 3491;
Best Local Similarity 50.3%; Pred. No. 1e-45;

Matches 159; Conservative 42; Mismatches 103; Indels 12; Gaps 5;

QY 1 VFVPPGGQPMGKRELLDASDVRESVACEAFAFYVDMSEVQLRSDPAAGLDYV 60
DB 57 VFVPPGGQPMGKRELLDASDVRESVACEAFAFYVDMSEVQLRSDPAAGLDYV 113
QY 61 DVVQPTLFAVMIISLAALMRSGVEPCAVLGHSLGEIAAAVSGGLSLADAAVVTLSQA 120
DB 114 EVVQPALFAVQTSIALALMRSGVEPCAVLGHSLGEIAAAVSGGLSLADAAVVTLSQA 173
QY 121 OTTLAGCALVSVAATPELLPRTAPWEDNPARLAANAANGPSTVSGAREAVDLVA 180
DB 174 MIPLVGMDMAVLSADEIEPRRIARMDD---VVLGVGPPSVLTGSPPEAVRVQ 229
QY 181 DLTAQVTRMIPVDVPAHSPLMTAIEERVVSGLLTPRPSRIPFHSSVGTGRDLREL 240
DB 230 ELSEAGRAQVYINSMASHAQNVDIAEGMSALAMPAGSEVPFVYSLTGAVDTREL 289
QY 241 DAAVYRNMSSTVREPARLLLOQPKTFVEMSPHPVLTGLOELAPDLGDTTGTADTV 300
DB 290 VADYWRSSPRLPVRPEDEIRALFVGPGTFEASBPVLAALQOTL---DAEGSSAAV 345
QY 301 IMGLRRGQGLDHL 316
DB 346 V-PTLGRGQGMKRP 360

RESULT 5

S43048

polyketide synthase type I - Streptomyces antibioticus
N/Contains: acyl carrier protein; acyltransferase; ketosynthase; ketoreductase; thio
C/Species: Streptomyces antibioticus
C/Date: 13-Jan-1995 #sequence_revision 06-Dec-1996 #text_change 26-May-2000
C/Accession: S43048; S41729
R/Swan, D.G.; Rodriguez, A.M.; Vilchee, C.; Mendez, C.; Salas, J.A.
Submitted to the EMBL Data Library, February 1993
A/Reference number: S43048
A/Molecule type: DNA
A/Residues: 1-3519 <SWA>
A/Cross-references: EMBL:L09654; NID:9153407; PIDN:AAA19695.1; PID:9153408
R/Swan, D.G.; Rodriguez, A.M.; Vilchee, C.; Mendez, C.; Salas, J.A.
Mol. Gen. Genet. 242, 358-362, 1994
A/Title: Characterisation of a Streptomyces antibioticus gene encoding a type I polyketi

A/Reference number: S41729; MUID:94150470; PMID:8107683
A/Accession: S41729
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1683-3238,3273-3303,'K',3305-3407,'T',3409-3462,'Y',3464-3516,'E',3518-3519
A/Cross-references: EMBL:L09654
C/Genetics:
A/Start codon: GTG
C/Superfamily: [acyl-carrier-protein] S-malonyltransferase homology; 3-oxoacyl-[acyl-ca
ogy; short-chain alcohol dehydrogenase homology
C/Keywords: antibiotic biosynthesis; carrier protein
F/57-462/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F/570-821/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F/1402-1381/Domain: short-chain alcohol dehydrogenase homology <SAD1>
F/1489-1560/Domain: acyl carrier protein homology <ACP1>
F/1708-2111/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F/2221-2502/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F/2858-3037/Domain: [acyl-carrier-protein] S-malonyltransferase homology <SAD2>
F/3143-3214/Domain: acyl carrier protein homology <ACP2>
F/3305-3500/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>

Query Match 43.7%; Score 707; DB 2; Length 3519;
Best Local Similarity 49.4%; Pred. No. 1.4e-44;

Matches 157; Conservative 41; Mismatches 106; Indels 14; Gaps 4;

QY 1 VFVPPGGQPMGKRELLDASDVRESVACEAFAFYVDMSEVQLRSDPAAGLDYV 60
DB 2221 VFVPPGGQPMGKRELLDASDVRESVACEAFAFYVDMSEVQLRSDPAAGLDYV 2279
QY 61 DVVQPTLFAVMIISLAALMRSGVEPCAVLGHSLGEIAAAVSGGLSLADAAVVTLSQA 120
DB 2280 DVVQPALFAVQTSIALALMRSGVEPCAVLGHSLGEIAAAVSGGLSLADAAVVTLSQA 2339
QY 121 OTTLAGCALVSVAATPELLPRTAPWEDNPARLAANAANGPSTVSGAREAVDLVA 178
DB 2340 IARIAGGGMVSVS-----LPAGRVRTMLDTYGRVSAVAVGSPSTVSGVQALDEL 2393
QY 179 VADLTAAQVTRMIPVDVPAHSPLMTAIEERVVSGLLTPRPSRIPFHSSVGTGRDLREL 236
DB 2394 LAGGERGVYARRVVDVYASHAQMOLRELEALADITPODSVFFSTVTDWLDIT 2453
QY 239 ELDAVYRNMSSTVREPARLLLOQPKTFVEMSPHPVLTGLOELAPDLGDTTGTAD 298
DB 2454 ALDAGVFTNLRETVRQOAVEGLVAGMGAFVESCSPHPVLTGLOELAPDLGDTTGTAD 2508
QY 299 TVIMGTLRGQGLDHL 316
DB 2509 AAVLGSLRDEGGLRL 2526

RESULT 6

T03224

probable polyketide synthase module 4 - Streptomyces hygroscopicus
C/Species: Streptomyces hygroscopicus
C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 26-May-2000
C/Accession: T03224
R/Ruan, X.; Stassei, D.; Lax, S.; Katz, L.
Gene 203, 1-9, 1997
A/Title: A second type-I PKS gene cluster isolated from Streptomyces hygroscopicus ATCC
A/Reference number: Z14848; MUID:98085969; PMID:9426000
A/Accession: T03224
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1937 <RUA>
A/Cross-references: EMBL:AF007101; NID:92624946; PIDN:ACG38064.1; PID:92624951
A/Experimental source: ATCC 29253
C/Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoa
homology: [acyl-carrier-protein] S-malonyltransferase homology
C/Keywords: carrier protein
F/55-456/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F/559-827/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
Query Match 43.2%; Score 699; DB 2; Length 1937;

Best Local Similarity 48.7%; Pred. No. 2.6e-44;
Matches 153; Conservative 39; Mismatches 100; Indels 22; Gaps 4;

QY 1 VFVPPGQPGMPGKRELLDSDVFRSVCRAEAFAPYVMSVEQVLRSDPADAGLDV 60
D 549 VFVPPGQPGMPGKRELLDSDVFRSVCRAEAFAPYVMSVEQVLRSDPADAGLDV 608
QY 61 DVVQPTLFPAWVMSLAALMRSGVPCAVLGHSLEIAAAHVSGLSLADARVYTLMSQA 120
D 609 DVVQPTLFPAWVMSLAALMRSGVPCAVLGHSLEIAAAHVSGLSLADARVYTLMSQA 668
QY 121 QTTLAGTALVSAATPDELLPRIAPWTEENPARLAVAANGPSTVSGAREAVADIVA 180
D 669 LTLAGEGCVVVRSAVYENLRQW----EGRLSAVANGPEAVVSGVAALELL- 723
QY 181 DLTAQVTRMIPDVPAHSPMLMAIERVVSGLPTPRSRIRPHSSVTGGLDREL 240
D 724 --ATEDARRVAADVASHAVERIEKLTRTLTDVQPMTSRVFLFSTVERDWDITASM 780
QY 241 DAAVYRNMSSTVFEPAARLLLOQPKTFVEMSPHPLTWLOELAPDLGDTTGTADTV 300
D 781 DTGWTYRNLRTQTFWFDVAGVLSAGFVFEVSPHPLTLASVRESCF-----EAV 831
QY 301 IMGTLRGQGLDHL 314
D 832 VTGTLRR-----DH 840

RESULT 7

T17464
tifamycin polyketide synthase modules 4-6 - Amycolatopsis mediterranei
C/Species: Amycolatopsis mediterranei
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
C/Accession: T17464

R;Schupp, T.
Submitted to the EMBL Data Library, December 1997
A/Reference number: Z18802
A/Accession: T17464
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-5069 <SCH>
A/Cross-references: EMBL:AJ223012; NID:e1227119; PID:e1227121; PIDN:CAA11036.1
A/Experimental source: Strain LBG A3136
C/Superfamily: acyl carrier protein homology
C/Keywords: carrier protein
F;1631-1702/Domain: acyl carrier protein homology <ACPI>
F;3238-3309/Domain: acyl carrier protein homology <ACP2>
F;4939-5010/Domain: acyl carrier protein homology <ACP3>

Query Match 43.1%; Score 697.5; DB 2; Length 5069;
Best Local Similarity 48.7%; Pred. No. 1.1e-43;
Matches 155; Conservative 41; Mismatches 107; Indels 15; Gaps 5;

QY 1 VFVPPGQPGMPGKRELLDSDVFRSVCRAEAFAPYVMSVEQVLRSDPADAGLDV 60
D 570 VFVPPGQPGMPGKRELLDSDVFRSVCRAEAFAPYVMSVEQVLRSDPADAGLDV 627
QY 61 DVVQPTLFPAWVMSLAALMRSGVPCAVLGHSLEIAAAHVSGLSLADARVYTLMSQA 120
D 628 DVVQPTLFPAWVMSLAALMRSGVPCAVLGHSLEIAAAHVSGLSLADARVYTLMSQA 687
QY 121 QTTLAGTALVSAATPDELLPRIAPWTEENPARLAVAANGPSTVSGAREAVADIVA 179
D 688 IASLARGGMAASVLAEBEDATLEPVA---GRVRAANGPTSVIADADALDEAL 743
QY 180 ADLTAQVTRMIPDVPAHSPMLMAIERVVSGLPTPRSRIRPHSSVTGGLDREL 238
D 744 DALDDQGRIRIRVAVDVAASHTRHVEAARDALAEMLGIRQAQAPVPFYSVTGGMVEDAG 803
QY 239 ELDAAYVYRNMSSTVFEPAARLLLOQPKTFVEMSPHPLTWLOELAPDLGDTTGTAD 298
D 804 VLDGWTYRNLRRQVFGPAVALLEEGHRAVFEVSAHPVLPVQINELVD-----TE 856

QY 299 TVIMGTLRGQGLDHL 316
D 857 AVVTGTLRRDGGIRLL 874

RESULT 8

T17409
polyketide synthase type I - Streptomyces venezuelae
C/Species: Streptomyces venezuelae
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
C/Accession: T17409

R;Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
A/Title: A gene cluster for macrolide antibiotic biosynthesis in streptomyces venezuelae
A/Reference number: Z18773; MUID:98445333; PMID:9770448
A/Accession: T17409
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-4613 <XUE>
A/Cross-references: EMBL:AF079138; NID:G3808326; PID:G3800834; PIDN:AAC69329.1
C/Genetics:
A/Gene: pikAI
C/Superfamily: acyl carrier protein homology
C/Keywords: antibiotic biosynthesis; carrier protein
F;1010-1081/Domain: acyl carrier protein homology <ACP1>
F;2495-2566/Domain: acyl carrier protein homology <ACP2>
F;4407-4478/Domain: acyl carrier protein homology <ACP3>

Query Match 43.0%; Score 696.5; DB 2; Length 4613;
Best Local Similarity 49.2%; Pred. No. 1.2e-43;
Matches 156; Conservative 42; Mismatches 100; Indels 19; Gaps 5;

QY 2 FVPPGQPGMPGKRELLDSDVFRSVCRAEAFAPYVMSVEQVLRSDPADAGLDV 61
D 634 FVPPGQPGMPGKRELLDSDVFRSVCRAEAFAPYVMSVEQVLRSDPADAGLDV 633
QY 62 DVVQPTLFPAWVMSLAALMRSGVPCAVLGHSLEIAAAHVSGLSLADARVYTLMSQA- 120
D 694 DVVQPTLFPAWVMSLAALMRSGVPCAVLGHSLEIAAAHVSGLSLADARVYTLMSQA 753
QY 121 QTTLAGTALVSAATPDELLPRIAPWTEENPARLAVAANGPSTVSGAREAVADIVA 180
D 754 AAHLAGKGMSTLALSDAVALERLAGPD-----GLSAANGPTATVSGDPVQIEBLAR 808
QY 181 DLTAQVTRMIPDVPAHSPMLMAIERVVSGLPTPRSRIRPHSSVTGGLDREL 240
D 809 ACBDDGRRARVIVDVASHSRQVEITSELAEVLAGISPPAPVPFSTLSEGMITEPVL 868
QY 241 DAAVYRNMSSTVFEPAAR-LLLOQPKTFVEMSPHPLTWLOELAPDLGDTTGTAD 299
D 869 DGGWTYRNLRRQVFGPAVALLEEGHRAVFEVSAHPVLPVQINELVD-----GTVTGLA- 920
QY 300 VIMGTLRGQGLDHL 316
D 921 ----TLRRDGGQDRLV 933

RESULT 9

H70621
probable polyketide synthase - Mycobacterium tuberculosis (strain H37Rv)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C/Accession: H70621

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.;
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Sgares, R.; Sultom, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98255987; PMID:9634230
A/Accession: H70621
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA

A/Residues: 1-2126 <COL>
A/Cross-references: GB:Z65992; GB:AL123456; NID:93261718; PIDN:CAB06632.1; PID:q1839013
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: pke7
C/Superfamily: mycrocetic acid synthase, 3-oxoacyl-[acyl-carrier-protein] synthase I h
name homology; [acyl-carrier-protein] S-malonyltransferase homology
C/Keywords: carrier protein
F/56-456/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F/563-845/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F/1694-1875/Domain: short-chain alcohol dehydrogenase homology <SADH>
F/1976-2051/Domain: acyl carrier protein homology <ACP2>
Query Match 42.6%; Score 689.5; DB 2; Length 2126;
Best Local Similarity 47.1%; Pred. No. 1.5e-43;
Matches 148; Conservative 46; Mismatches 115; Indels 5; Gaps 2;
Qy 2 FVPPGQGFQWPMGRBELDASDVRESVACEAAPPYVDMSVQVLRDSDPAQLRDV 61
Db FVPPGQGFQWPMGRBELDASDVRESVACEAAPPYVDMSVQVLRDSDPAQLRDV 61
Qy 564 FVPPGQGFQWPMGRBELDASDVRESVACEAAPPYVDMSVQVLRDSDPAQLRDV 623
Db 564 FVPPGQGFQWPMGRBELDASDVRESVACEAAPPYVDMSVQVLRDSDPAQLRDV 623
Qy 62 VVQPTLPFVMTSLAALMPSQGVPCAVTGHSLGELIAAHVSGSLDAARVVTLMGQAQ 121
Db 624 VVQPTLPFVMTSLAALMPSQGVPCAVTGHSLGELIAAHVSGSLDAARVVTLMGQAQ 121
Qy 122 TLLAGTALVVAATPDELPRIPWTEEDNPRLAVALAANGPRSTVSGAREAVADLVAD 181
Db 684 VRLGAGAGMVSLLAGGQPPAEKLAGMGD---RLNIAAVNGSVSVTLAGEIDATYELMQR 739
Qy 182 LTAQVTRMTIPVDVPAHSPMLYALIEKRVSGLLPTPRSPRIPEHSSVTGRLDREL 241
Db 740 CEAGIRARRIDVDVASHAQVDALRELLAALRGIEPRSTVAFFSTVTELMDTAGVN 799
Qy 242 AAAYYRNMSSTVREPPARLLLOQPKTFVEMSPHPVTMLGELAPDLGTTGTADVI 301
Db 800 AEVYRSLRQPVQGERAARNAPFDGQVRFVSSPHPLVLAIGETLVDC-DRGATGEPIV 858
Qy 302 MGTLRGQGTLDHF 315
Db 859 IPTLRDQGVGRF 872
RESULT 10
S22012
erythronolide synthase (EC 2.3.1.94) III - Saccharopolyspora erythraea
N/Alternate names: 6-deoxyerythronolide B synthase III; protein A
C/Species: Saccharopolyspora erythraea
C/Date: 07-Apr-1994 #sequence revision 07-Apr-1994 #text_change 26-May-2000
C/Accession: S22012; S34784; S23206
R/Bevitt, D.J.
Submitted to the EMBL Data Library, September 1991
A/Reference number: S22011
A/Accession: S22012
A/Molecule type: DNA
A/Residues: 1-3172 <BEV1>
A/Cross-references: EMBL:X62569; NID:946977; PIDN:CAA44449.1; PID:g46979
R/Experimental source: strain NRRL 2338
R/Bevitt, D.J.; Cortes, J.; Haydock, S.F.; Leadlay, P.F.
Eur. J. Biochem. 204, 39-49, 1992
A/Title: 6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of
A/Reference number: S23070; MUID:923155230; PMID:1740151
A/Accession: S23070
A/Molecule type: DNA
A/Residues: 1-715, 717-728, 'C', '729-886, 888-894, 'QR', 895-910, 'E', 912-1276, 'RR', 1279-1284, 'EV2'
A/Cross-references: EMBL:X62569
R/Gaffrey, P.; Bevitt, D.J.; Staunton, J.; Leadlay, P.F.
FEBS Lett. 304, 225-228, 1992
A/Title: Identification of DEBS 1, DEBS 2 and DEBS 3, the multienzyme polypeptides of th
A/Accession: S23206
A/Molecule type: protein
A/Residues: 2-11, 'XXX', 15 <CAF>

A/Experimental source: strain CA340
C/Genetics:
A/Gene: eryA
C/Function:
A/Description: catalyzes the construction of a polyketide chain, which is then cyclised
A/Pathway: erythromycin biosynthesis
C/Superfamily: [acyl-carrier-protein] S-malonyltransferase homology; 3-oxoacyl-[acyl-ca
oyl] short-chain alcohol dehydrogenase homology
C/Keywords: acyltransferase; antibiotic biosynthesis; carrier protein; coenzyme A; mult
F/60-450/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F/555-835/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F/1117-1297/Domain: short-chain alcohol dehydrogenase homology <SAD1>
F/1392-1463/Domain: acyl carrier protein homology <ACP1>
F/1511-1914/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F/2022-2300/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F/2557-2734/Domain: short-chain alcohol dehydrogenase homology <SAD2>
F/2819-2890/Domain: acyl carrier protein homology <ACP2>
F/2859-3153/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACP3>
Query Match 42.6%; Score 689; DB 2; Length 3172;
Best Local Similarity 46.8%; Pred. No. 2.7e-43;
Matches 148; Conservative 52; Mismatches 102; Indels 14; Gaps 5;
Qy 3 VFPQGFQWPMGRBELDASDVRESVACEAAPPYVDMSVQVLRDSDPAQLRDV 62
Db 557 VFPQGFQWPMGRBELDASDVRESVACEAAPPYVDMSVQVLRDSDPAQLRDV 613
Qy 63 VQPTLPFVMTSLAALMPSQGVPCAVTGHSLGELIAAHVSGSLDAARVVTLMGQAQ 122
Db 614 VQPTLPFVMTSLAALMPSQGVPCAVTGHSLGELIAAHVSGSLDAARVVTLMGQAQ 122
Qy 123 TLLAGTALVVAATPDELPRIPWTEEDNPRLAVALAANGPRSTVSGAREAVADLVAD 181
Db 674 RLGGGGMABFGCTEBAARIGRA---GALSTASVNGPRSVVAVAGSGPLDELIAE 729
Qy 182 LTAQVTRMTIPVDVPAHSPMLYALIEKRVSGLLPTPRSPRIPEHSSVTGRLDREL 241
Db 730 CEAAHARRIPVDVASHSPQVSLRELLTLAGISPVADVALVSTTGQPIDTATMD 789
Qy 242 AAAYYRNMSSTVREPPARLLLOQPKTFVEMSPHPVTMLGELAPDLGTTGTADVI 301
Db 790 TAYVYANLRQVRFQDTRQLAAGFPAFVSPHPVLTIVGIEATL---DSALPADGA 845
Qy 300 VIMGTLRGQGTLDHF 315
Db 846 CVVGTLRDRGGLADF 861
RESULT 11
S13595
6-deoxyerythronolide B synthase - Saccharopolyspora erythraea
N/Alternate names: protein A
C/Species: Saccharopolyspora erythraea
C/Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 26-May-2000
C/Accession: S13595
R/Cortes, J.; Haydock, S.F.; Roberts, G.A.; Bevitt, D.J.; Leadlay, P.F.
Nature 348, 176-178, 1990
A/Title: An unusually large multifunctional polypeptide in the erythromycin-producing E
A/Reference number: S13595; MUID:91043075; PMID:2234082
A/Accession: S13595
A/Molecule type: DNA
A/Residues: 1-3178 <COR>
A/Cross-references: EMBL:X6107; NID:946975; PIDN:CAA39503.1; PID:g46976
A/Experimental source: strain NRRL 2338
C/Genetics:
A/Gene: eryA
C/Function:
A/Description: catalyzes the construction of a polyketide chain, which is then cyclised
A/Pathway: erythromycin biosynthesis
C/Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier prote
-carrier-protein] S-malonyltransferase homology
C/Keywords: acyltransferase; antibiotic biosynthesis; carrier protein; multifunctional
F/60-450/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>

F:553-833/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
 F:1133-1303/Domain: short-chain alcohol dehydrogenase homology <SAD1>
 F:1138-1469/Domain: acyl carrier protein homology <ACP1>
 F:1517-1920/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
 F:2028-2306/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
 F:2553-2740/Domain: short-chain alcohol dehydrogenase homology <SAD2>
 F:2853-2896/Domain: acyl carrier protein homology <ACP2>
 F:2965-3159/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>

Query Match 42.6%; Score 689; DB 2; Length 3178;
 Best Local Similarity 46.8%; Pred. No. 2.8e-43;
 Matches 148; Conservative 52; Mismatches 102; Indels 14; Gaps 5;

QY 3 VFPGGQPMQMGRELIDASDVFRSVRACEAFAFYDMSVEQVLRDSDPAPGLDRVD 62
 DB 555 VFPGGAGMOGMARLDLSEQVFDASIRDCERALAPHDMSITDILL--SGARPLDRVD 611
 QY 63 VQPTLFAVMISLAALMRSQGVPCAVLGHSLGEIAAAHVSGGLSLADAAVVTLMQAQT 122
 DB 612 VQPALFAVMISLAALMRSQGVPCAVLGHSLGEIAAAHVSGGLSLADAAVVTLMQAQT 671
 QY 123 TLATGALVSAATDELPRIPAPTEENPARLAANAANGPRS-TVSGAREAAVDLYAD 181
 DB 672 RLGGGGMASFGGLGTQQAERIGRPA---GALSLASVNGPRSVVVAAGESGPDDELIAE 727
 QY 182 LTAQVTRMTIPVDVPAHSPLMTAIEERVSGGLPTPRSPRIPFHSSVTGRLDTRLD 241
 DB 728 CEAEHKKARRIPVDYASHSPVESIREBELTELAGISVSADVALYSTTTQPIDTATMD 787
 QY 242 AAYWRNMSSTVFEPAARLLLOQPKTFVEMSPHVLTMGLQELAPDLGDTTGTAD--T 299
 DB 788 TAYVANLRQVRFQDAFROLAAGFADFVEVSPHVLTVGIEATL----DSALPADAGA 843
 QY 300 VIMGTLRRCGQTLDFH 315
 DB 844 CVGTLRRDRGGLADF 859

RESULT 12

T17410
 polyketide synthase type I - Streptomyces venezuelae
 C/Species: Streptomyces venezuelae
 C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
 C/Accession: T17410

R:Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
 Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998

A/Title: A gene cluster for macrocyclic antibiotic biosynthesis in streptomyces venezuelae
 A/Reference number: Z18773; MUID:98445333; PMID:9770448
 A/Accession: T17410

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
 A/Residues: 1-3739 <XUE>
 A/Cross-references: EMBL:AF079138; NID:g3808326; PID:g3800835; PIDN:AAC69330.1
 C/Genetics:

A/Gene: pikAII
 C/Superfamily: acyl carrier protein homology
 C/Keywords: antibiotic biosynthesis; carrier protein
 F:1445-1516/Domain: acyl carrier protein homology <ACP1>
 F:3570-3641/Domain: acyl carrier protein homology <ACP2>

Query Match 42.1%; Score 681.5; DB 2; Length 3739;
 Best Local Similarity 46.9%; Pred. No. 1.2e-42;
 Matches 152; Conservative 42; Mismatches 98; Indels 19; Gaps 5;

QY 2 FVFGGQPMQMGRELIDASDVFRSVRACEAFAFYDMSVEQVLRDSDPAPGLDRVD 61
 DB 2069 FVFGGQPMQMGRELIDASDVFRSVRACEAFAFYDMSVEQVLRDSDPAPGLDRVD 2128
 QY 62 VQPTLFAVMISLAALMRSQGVPCAVLGHSLGEIAAAHVSGGLSLADAAVVTLMQA- 120
 DB 2129 VQPTLFAVMISLAALMRSQGVPCAVLGHSLGEIAAAHVSGGLSLADAAVVTLMQA- 2188
 QY 121 OTTAGTALVSAATDELPRIPAPTEENPARLAANAANGPRS-TVSGAREAAVDLYA 180

DB 2189 GAHLAAGGCMISLASBAVERLAGFD-----GLSVAAVNGPFAVVSQDPQIQELAQ 2243

QY 181 DLTAQVTRMTIPVDVPAHSPLMTAIEERVSGGLPTPRSPRIPFHSSVTGRLDTRLD 240

DB 2244 ACEADGVRARIPVDYASHSPVESIREBELTELADVLGASFPQTPQVFPFSTLEGAWITPEAL 2303

QY 241 DAAVWRNMSSTVFEPAAR-LLLOQPKTFVEMSPHVLTMGLQELAPDLGDTTGTADT 299

DB 2304 DGTWYNLNRHVRGFAPAVETLARDSEGFTHFEVSAHPVLTMLPE-----TVTG----- 2353

QY 300 VIMGTLRRCGQ 310

DB 2354 --LGTLRDRNG 2362

RESULT 13

T17411
 polyketide synthase III - Streptomyces venezuelae
 C/Species: Streptomyces venezuelae
 C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 03-Nov-2000
 C/Accession: T17411

R:Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
 Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998

A/Title: A gene cluster for macrocyclic antibiotic biosynthesis in streptomyces venezuelae
 A/Reference number: Z18773; MUID:98445333; PMID:9770448
 A/Accession: T17411

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA
 A/Residues: 1-1562 <XUE>

A/Cross-references: EMBL:AF079138; NID:g3808326; PID:g3800836; PIDN:AAC69331.1
 C/Genetics:

A/Gene: pikAIII
 C/Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoacyl
 homology; [acyl-carrier-protein] S-malonyltransferase homology
 C/Keywords: antibiotic biosynthesis; carrier protein
 F:1403-1474/Domain: acyl carrier protein homology <ACP>

Query Match 41.7%; Score 674; DB 2; Length 1562;
 Best Local Similarity 48.1%; Pred. No. 1.5e-42;
 Matches 149; Conservative 41; Mismatches 102; Indels 18; Gaps 4;

QY 2 FVFGGQPMQMGRELIDASDVFRSVRACEAFAFYDMSVEQVLRDSDPAPGLDRVD 61

DB 565 FVFGGQPMQMGRELIDSSPERFASMACTATLSYVMSLEAVVRQBPAPFTDTRVD 624

QY 62 VQPTLFAVMISLAALMRSQGVPCAVLGHSLGEIAAAHVSGGLSLADAAVVTLMQA- 120

DB 625 VQPTLFAVMISLAALMRSQGVPCAVLGHSLGEIAAAHVSGGLSLADAAVVTLMQA- 684

QY 121 OTTAGTALVSAATDELPRIPAPTEENPARLAANAANGPRS-TVSGAREAAVDLYA 180

DB 685 AAHLAAGGCMISLASBAVERLAGFD-----GLSVAAVNGPFAVVSQDPQIQELAR 739

QY 181 DLTAQVTRMTIPVDVPAHSPLMTAIEERVSGGLPTPRSPRIPFHSSVTGRLDTRLD 240

DB 740 TCERADGVRARIPVDYASHSPVESIREBELTELADVLGASFPQTPQVFPFSTLEGAWITPEAL 799

QY 241 DAAVWRNMSSTVFEPAARLLLOQPKTFVEMSPHVLTMGLQELAPDLGDTTGTADTV 300

DB 800 DGTWYNLNRHVRGFAPAVETLARDSEGFTHFEVSAHPVLTMLPE-----TVTG----- 848

QY 301 VIMGTLRRCGQ 310

DB 849 --LGTLRREG 857

RESULT 14

T17465
 rifamycin polyketide synthase module 7 - Amycolatopsis mediterranei
 C/Species: Amycolatopsis mediterranei
 C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 03-Nov-2000
 C/Accession: T17465

R:Schupp, T.
 submitted to the EMBL Data Library, December 1997
 A:Reference number: Z18802
 A:Accession: T17465
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1763 <SCN>
 A:Cross-references: EMBL:AJ223012; NID:e1227119; PID:e1227122; PIDN:CAA11037.1
 A:Experimental source: strain LBG A3136
 C:Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoad
 homology; [acyl-carrier-protein] 8-malonyltransferase homology
 C:Keywords: carrier protein
 F:15-1686/Domain: acyl carrier protein homology <ACP>

Query Match 41.6%; Score 673.5; DB 2; Length 1763;
 Best Local Similarity 47.8%; Pred. No. 1.9e-42;
 Matches 152; Conservative 39; Mismatches 108; Indels 19; Gaps 5;

Qy 1 VFVPGGQPGWPGMGRELLDASDVRESVRACEAFAFYDMSVEOVLRLSDPAPGLDRV 60
 Db 555 VVVFPGGQPGWPGMGRELLDASDVRESVRACEAFAFYDMSVEOVLRLSDPAPGLDRV 611
 Qy 61 DVVPTLPFAVMISLAALMRSGVPCAVLGHSLGEIAAAHVSGLSLADARVYTLMSQA 120
 Db 612 EVLQPCFPAVWVGLAAVWESAGVPRDAVGHSGEIAAASVSGALTLDDAKVVALRSQA 671
 Qy 121 -QTLTLAGTALVSAATPDELPRIPATWEDNPARLVAANVGPRSTVSGAREVADLV 179
 Db 672 IAAHLSGRGMAVSLASDEANARLGLW---DGRIEVAANVGPRSTVSGAREVADLV 727
 Qy 180 ADLTAQVTRTMIIPDVPAHSPMLVYAEERVSGLLPTPRPSRIPFHSSVTGRL-DTR 238
 Db 728 EVLAGDGRVAVQVADVASHRVEDIRDTLAETLAGITTAQAPVPRFSTVTGGVNRDAD 787
 Qy 239 ELDAAYTRNMSSTVRFEPAR-LILQOGPKTFVEMSPHPVLTMLQELAPPLGDTTGTAD 298
 Db 788 VLDGQWYRNLRNQVRFPAPVAVLELQGHGVFEVSAPVLPQPSLEL-----TD 837
 Qy 299 TVIMGTLRRGGTLDFHL 316
 Db 838 AVVTGTLRRDDGRLRL 855

RESULT 15
 T17412
 polyketide synthase IV - Streptomyces venezuelae
 C:Species: Streptomyces venezuelae
 C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
 C:Accession: T17412
 R:Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
 Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
 A:Title: A gene cluster for macrolide antibiotic biosynthesis in streptomyces venezuelae
 A:Reference number: Z18773; MUID:98445333; PMID:9770448
 A:Accession: T17412
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1346 <XUB>
 A:Cross-references: EMBL:AF079138; NID:g3808326; PID:g380837; PIDN:AAC69332.1
 C:Gene: C1
 A:Gene: pikaiv
 C:Superfamily: acyl carrier protein homology
 C:Keywords: antibiotic biosynthesis; carrier protein
 F:1945-1016/Domain: acyl carrier protein homology <ACP>

Query Match 41.3%; Score 667.5; DB 2; Length 1346;
 Best Local Similarity 47.6%; Pred. No. 3.9e-42;
 Matches 148; Conservative 45; Mismatches 99; Indels 19; Gaps 6;

Qy 2 FVPPGGQPGWPGMGRELLDASDVRESVRACEAFAFYDMSVEOVLRLSDPAPGLDRV 61
 Db 562 FVPPGGQPGWPGMGRELLDASDVRESVRACEAFAFYDMSVEOVLRLSDPAPGLDRV 621
 Qy 62 VVQPTLPFAVMISLAALMRSGVPCAVLGHSLGEIAAAHVSGLSLADARVYTLMSQA- 120

Db 622 VVQPTLPFAVMISLAALMRSGVPCAVLGHSLGEIAAAHVSGLSLADARVYTLMSQA 681
 Qy 121 QTLTLAGTALVSAATPDELPRIPATWEDNPARLVAANVGPRSTVSGAREVADLV 180
 Db 682 AAHLAGRGMTISLASEATRORI---SNLHGLSLAAVNGPFAVVSQDPQIOELAQ 736
 Qy 181 DLTAQVTRTMIIPDVPAHSPMLVYAEERVSGLLPTPRPSRIPFHSSVTGRLDTRRL 240
 Db 737 ACBAGDIRAIIIPDVASHASHAHTIENELADVLASLPQTPQVPRFSTVTGGVNRDAD 796
 Qy 241 DAAVYTRNMSSTVRFEPAR-LILQOGPKTFVEMSPHPVLTMLQELAPPLGDTTGTAD 299
 Db 797 DCGWYRNLRNQVRFPAPVAVLELQGHGVFEVSAPVLPQPSLEL-----PD--KVTGLA-- 848
 Qy 300 VIMGTLRRGG 310
 Db 849 ---TLRRDDG 855

Search completed: June 17, 2003, 13:12:19
 Job time : 8.05626 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 13:08:02 ; Search time 9.84594 Seconds
(without alignments)
3431.399 Million cell updates/sec

Title: US-09-914-286-3_COPY_29_344

Perfect score: 1618
Sequence: 1 VFVPCGQPGWPGWRELD.....ADTVIMGTLRGQGLDHL 316

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubppa/PTCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubppa/PTCTUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	819	50.6	344	9 US-09-735-056-33	Sequence 33, Appl
2	714.5	44.2	3816	9 US-09-808-880-3	Sequence 3, Appl
3	712	44.0	4150	9 US-09-808-880-2	Sequence 2, Appl
4	707	43.7	3519	9 US-09-808-880-4	Sequence 4, Appl
5	696.5	43.0	4551	9 US-09-793-708-1	Sequence 1, Appl
6	696.5	43.0	4613	9 US-09-860-846-31	Sequence 31, Appl
7	696.5	43.0	4613	9 US-09-988-384B-31	Sequence 31, Appl
8	696.5	43.0	4613	9 US-09-836-821-31	Sequence 31, Appl
9	696.5	43.0	4613	10 US-09-861-289-31	Sequence 31, Appl
10	696.5	43.0	11877	9 US-09-860-846-6	Sequence 6, Appl
11	696.5	43.0	11877	9 US-09-836-821-6	Sequence 6, Appl
12	696.5	43.0	11877	10 US-09-861-289-6	Sequence 6, Appl
13	696.5	43.0	12199	9 US-09-988-384B-6	Sequence 6, Appl
14	696.5	43.0	3739	9 US-09-860-846-33	Sequence 33, Appl
15	696.5	43.0	3739	9 US-09-988-384B-33	Sequence 33, Appl
16	696.5	43.0	3739	9 US-09-836-821-33	Sequence 33, Appl
17	696.5	43.0	3739	9 US-09-793-708-2	Sequence 2, Appl
18	696.5	43.0	3739	10 US-09-861-289-33	Sequence 33, Appl
19	674	41.7	1562	9 US-09-860-846-35	Sequence 35, Appl

20	674	41.7	1562	9 US-09-888-384B-35	Sequence 35, Appl
21	674	41.7	1562	9 US-09-836-821-35	Sequence 35, Appl
22	674	41.7	1562	9 US-09-793-708-3	Sequence 3, Appl
23	674	41.7	1562	10 US-09-861-289-35	Sequence 35, Appl
24	667.5	41.3	1346	9 US-09-860-846-37	Sequence 37, Appl
25	667.5	41.3	1346	9 US-09-988-384B-37	Sequence 37, Appl
26	667.5	41.3	1346	9 US-09-836-821-37	Sequence 37, Appl
27	667.5	41.3	1346	9 US-09-793-708-37	Sequence 37, Appl
28	667.5	41.3	1346	10 US-09-861-289-37	Sequence 37, Appl
29	667	41.2	345	9 US-09-735-056-34	Sequence 34, Appl
30	627.5	38.8	3798	9 US-10-014-717-6	Sequence 6, Appl
31	583.5	36.1	7257	9 US-10-014-717-5	Sequence 5, Appl
32	574.5	35.5	1832	9 US-10-014-717-4	Sequence 4, Appl
33	441.5	27.3	1837	9 US-09-712-363-261	Sequence 261, App
34	411	25.4	2439	9 US-10-014-717-7	Sequence 7, Appl
35	404.5	25.0	343	9 US-09-735-056-32	Sequence 32, Appl
36	399.5	24.7	1616	9 US-09-712-363-262	Sequence 262, App
37	398	24.6	1421	9 US-10-014-717-2	Sequence 2, Appl
38	395	24.4	5215	9 US-09-860-846-2	Sequence 2, Appl
39	395	24.4	5215	9 US-09-988-384B-2	Sequence 2, Appl
40	395	24.4	5215	9 US-09-836-821-2	Sequence 2, Appl
41	395	24.4	5215	10 US-09-861-289-2	Sequence 2, Appl
42	386.5	23.9	1402	9 US-09-712-363-166	Sequence 166, App
43	383	23.7	328	9 US-09-735-056-31	Sequence 31, Appl
44	310	19.2	1610	9 US-09-738-626-6666	Sequence 6666, Ap
45	274.5	17.0	2563	9 US-09-836-705-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-09-735-056-33
Sequence 33, Application US/09735056
Publication No. US20030013662A1
GENERAL INFORMATION:
APPLICANT: Katz, Leonard
APPLICANT: Stassi, Diane L.
APPLICANT: Sumner Jr., Richard G.
APPLICANT: Ruan, Xiaoen
APPLICANT: Pereda-Lopez, Ana
APPLICANT: Kakava, Stephan J.
TITLE OF INVENTION: NOVEL POLYPEPTIDE DERIVATIVES
TITLE OF INVENTION: AND RECOMBINANT METHODS FOR MAKING SAME
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Rd.
CITY: Abbott Park
STATE: Illinois
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/735, 056
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/856,003
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dianne Caetano
REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 4952.US.P2
TELEPHONE: (847)-938-3137
TELEFAX: (847)-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. US20030013662A1e
US-09-735-056-33

Query Match 50.6%; Score 819; DB 9; Length 344;
Best Local Similarity 54.0%; Pred. No. 4.2e-61;
Matches 168; Conservative 38; Mismatches 95; Indels 10; Gaps 3;

QY 1 VFVPPGQGPQWPGMGRELLDASDVFRSEVRAACEAFAFYDMSVEQVLRDSPDAPGLDRV 60
DB 5 VFVPPGQGSQWAGMAEGLLERSGAFRSADSCDALRRLYGLMSVLSVLRGPDAPSLDRV 64
QY 61 DVQOPTLFAMVLSLAALMRSGVEPCAVLHSGEIAAAHVSGLSLADARVYVTLMSQA 120
DB 65 DVQOPVLFYMWVLSLAALMRSGVEPCAVLHSGEIAAAHVSGLSLADARVYVTLMSQA 124
QY 121 QTTLAGTALVSVAATPDELLP-RIAPWTEBNPARLAVAANGPSTVVSAGREAVADLVA 180
DB 125 WGLAGAGGMAVAPMPABELRLVLTWGD---RLAFAAASPGSCAVAGDPALAEALVA 180
QY 181 DLTAAYVTRMIP-VDVPAHSPLMYALIERVVSGLPTPRPSRIPFHSSVTGRLDTR 239
DB 181 LLTGEVHARPIPGVDTAGHSPOVDALRAHLLEVLAFAVAPRPADIPFSTVTGGLDSTE 240
QY 240 LDAAYVTRMSTVREPEARLLLOQGPRTVEMSPHVLTMGLQELAPDGLDGTGADT 299
DB 241 LDAAYVTRMSTVREPEARLLLOQGPRTVEMSPHVLTMGLQELAPDGLDGTGADT 299
QY 300 VIMGTLRRGQG 310
DB 296 AVLGTLRRRHG 306

RESULT 2
US-09-808-880-3
Sequence 3, Application US/09808880
Publication No. US20030027287A1
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT FILING DATE: US/09/808,880
CURRENT APPLICATION NUMBER: 2001-03-14
PRIOR FILING DATE: 1999-10-28
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/120,254
PRIOR FILING DATE: 1999-02-16
PRIOR APPLICATION NUMBER: 60/106,100
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 3816
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: Oleandolide PKS
US-09-808-880-3

Query Match 44.2%; Score 714.5; DB 9; Length 3816;
Best Local Similarity 50.3%; Pred. No. 6.1e-51;
Matches 160; Conservative 37; Mismatches 106; Indels 15; Gaps 4;
QY 1 VFVPPGQGPQWPGMGRELLDASDVFRSEVRAACEAFAFYDMSVEQVLRDSPDAPGLDRV 60

DB 567 VFVPPGQGSQWAGMAAGLLDACPFAEVAECAVADPLTGWSLVEVLRGGEAV--LGRV 624
QY 61 DVQOPTLFAMVLSLAALMRSGVEPCAVLHSGEIAAAHVSGLSLADARVYVTLMSQA 120
DB 625 DVQOPALMAVMSLARWRYYGVEPAVVGHSOGELIAACVAGSLADGARVYVLRSA 684
QY 121 QTTLAGTALVSVAATPDELLP--RIAPWTEBNPARLAVAANGPSTVVSAGREAVADL 178
DB 685 IARIAGGGWVS-----LPAGRVYTMDDTYGRLSVAAVNGPSTVVSAGDQALDEL 738
QY 179 VADITTAQVTRMIPVDVPAHSPLMYALIERVVSGLPTPRPSRIPFHSSVTGRLDTR 238
DB 739 LAGGERGVARRVPVDYASHSQMDQLRDLEALADITPHSHSVFFSTVADWLDTT 798
QY 239 ELDAAYVTRMSTVREPEARLLLOQGPRTVEMSPHVLTMGLQELAPDGLDGTGAD 298
DB 799 ALDGGVFTMLRETFRQAEVBEGLVAQMGAFPCSHPLVVGIEQTL-----DIYEAD 853
QY 299 TVIMGTLRRGQGLDHEL 316
DB 854 AVALGSLRDEGGIGRL 871

RESULT 3
US-09-808-880-2
Sequence 2, Application US/09808880
Publication No. US20030027287A1
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT FILING DATE: US/09/808,880
CURRENT APPLICATION NUMBER: 2001-03-14
PRIOR FILING DATE: 1999-10-28
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/120,254
PRIOR FILING DATE: 1999-02-16
PRIOR APPLICATION NUMBER: 60/106,100
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4150
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: Oleandolide PKS
US-09-808-880-2

Query Match 44.0%; Score 712; DB 9; Length 4150;
Best Local Similarity 50.3%; Pred. No. 1.1e-50;
Matches 160; Conservative 37; Mismatches 107; Indels 14; Gaps 4;

QY 1 VFVPPGQGPQWPGMGRELLDASDVFRSEVRAACEAFAFYDMSVEQVLRDSPDAPGLDRV 60
DB 1581 VFVPPGQGSQWAGMAAGLLDACPFAEVAECAVADPLTGWSLVEVLRGGEAV--LGRV 1639
QY 61 DVQOPTLFAMVLSLAALMRSGVEPCAVLHSGEIAAAHVSGLSLADARVYVTLMSQA 120
DB 1640 DVQOPALMAVMSLARWRYYGVEPAVVGHSOGELIAACVAGSLADGARVYVLRSA 1699
QY 121 QTTLAGTALVSVAATPDELLP--RIAPWTEBNPARLAVAANGPSTVVSAGREAVADL 178
DB 1700 IARIAGGGWVS-----LPAGRVYTMLEEFGRVSVAAVNGPSTVVSAGDQALDEL 1753
QY 179 VADITTAQVTRMIPVDVPAHSPLMYALIERVVSGLPTPRPSRIPFHSSVTGRLDTR 238
DB 1754 LAGGERGVARRVPVDYASHSQMDQLRDLEALADITPHSHSVFFSTVADWLDTT 1813

QY 239 ELDAAYVYRNNSSTVRFPPAARLLIQGPPTVEVMSPPHVLTMGQELAPLDIGDPTTGAD 298
 Db 1814 ALDAGYVFTNLRETIVRFQEAIVEGLVAQMGAFVESCSPHVLVPGITETL-----DTPAD 1866
 QY 299 TVIMGTLRQCGTLDHFL 316
 Db 1869 AVALSSLRDEGGDLRFL 1886

RESULT 4
US-09-80

```

Sequence 4, Application US/09808880
Publication No. US20030027287A1
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniell, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/808, 880
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: US/09/428, 517
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/120,254
PRIOR FILING DATE: 1999-02-16
PRIOR APPLICATION NUMBER: 60/106,100
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 3519
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-808-880-4

```

Query Match	Score	DB	Length
43.7%	707	9	3519

Best Local Similarity 49.4%; Pred. NO. 2.3e-50;
Matches 157; Conservative 41; Mismatches 106; Indels 14; Gaps 4;

Qy	1	VFVPPGGQPPMPGMRRLDSDVYRSEVRACEAFAYVYDMSVQVLRDSDAGGLDGRV	60
Db	2221	VFVPPGGQSPVGNAAAGLLDACPVAFAEACAAALDVTGMSLVEVQ-GRDAIVLGRV	22779
Qy	61	DVQVPTLFAMVLSLAALMRSGQVEPCAVLHSGEITAAAHVSGGLSLADAPRVYTLMSQA	120
Db	2280	DVQVPAALAAVMSILARTMYRGVEPAAVVGHSGEITAAACVAGGISLADGARVYVLRISA	23393
Qy	121	QTTLAGTAGALVSVAATPDELLP--RIAPMTEDNPRLVAANGRGSTVYSGAREAVDL	178
Db	2340	IARLAGGGGMSV-----LPAGIVRMTLDITGGRVSAANGSGSTVYSGDVALDEL	2399
Qy	179	VADLTAAQVTRMTIPVDVPAHSPLMYAIEERVVSGGLPITPPPSRIPIHSSVYTGRLDTR	238
Db	2394	LAGEGREGVRRARVVDVASHSAQMDQRLDELLLEALDITPDSSVFPFSTVYTDWMLDPTT	24533
Qy	239	ELDAAVYVRNNSSTYRFEPARLLTLOQPKTFVEMSPRPVYTMGQELAPDLGDTTAD	298
Db	2454	ALDAGYWTNLRETRYRPEAVBGLVAQMGAFVBCSPHPVLVPGIEQLDALDQ-----N	2508
Qy	299	TVIMGTLRRGGGLDHL 316	
Db	2509	AAVGLSLRDEGGDLRLL 2526	

RESULT 5
US-09-793-708-1

```

; Sequence 1, Application US/09793708
; Publication No. US20030104597A1
; GENERAL INFORMATION:

```

```

APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: McDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002121
CURRENT APPLICATION NUMBER: US/09/793,708
CURRENT FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: US 09/657,440
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: US 09/320,878
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: US 09/141,908
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: US 09/073,538
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: US 08/846,247
PRIOR FILING DATE: 1997-04-30
PRIOR APPLICATION NUMBER: US 60/134,990
PRIOR FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 4551
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-793-708-1

```

Query Match	43.0%	Score 696.5	DB 9	Length 4551
Best Local Similarity	49.2%	Pred. No. 2.5e-49		
Matches 156, Conservative	42	Mismatches 100	Indels 19	Gaps 5

```

Qy      2 FVPPGGGPPMPGAGELLDDASVPFESVRACEAAPAPVMDSVQVLRDSDPAGLDRVD 61
        |||||
Db      572 FVPPGGGTGMAAGAGELDDSSAVFPAANAACECLASLPVMDSVLAVVRQAAGATLERVD 631
        |||||

Qy      62 VVQPTFAVMI SLAALMRSQGVPCAVLGHS LGELIAAAHVS GGSGLADDAARVVTLMASCA- 120
        |||||
Db      632 VVQPTFAVMSVLARVQHGHGVTTPQAVVGHSGGELIAAAVYAGALS LDDDAARVVTLRKSI 699
        |||||

Qy      121 QTTLAGTGA LSVAAITPELLPR IAPWTE DNPARLA VAAVNGPSTVSGAREAVDLVA 180
        |||||
Db      692 AALHLAGKGMLSLALSEDAVLERLNGFD-----GLSVAAVNGPFAVTVSGDPVQIBELAR 746
        |||||

Qy      181 DLTLAAVTRTMI PVDVPAHSPMLMAIIEERVVSGLLPTTPRPSRI PFHSSVTGRLDTRREL 240
        |||||
Db      747 ACEADQVRA RAVIPVYASHSRQVEIIESELAEVLGAS PQAPRPVPFSTLEGAWITEPVL 806
        |||||

Qy      241 DAAVYWRMNSSTVYREPPAAR-LLDQGGKTIVENSPPHVLTMGQELAPDLGGDTTGTGADT 299
        |||||
Db      807 DGGYWRNRLRH RVGAPAVETLATDEGTHFEVVSAPHLTMALP-----GTVTGLA-- 858
        |||||

Qy      300 VIMGTLRGGQGLDHL 316
        |||||
Db      859 ---TLRDNGGQDRLV 871
        |||||

```

RESULT 6
US-09-860-846-31

Sequence 31, Application US/09860846
Patent No. US2002016742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D. H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438U1
CURRENT APPLICATION NUMBER: US/09/860,846
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26

NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-31

Query Match 43.0%; Score 696.5; DB 9; Length 4613;
Best Local Similarity 49.2%; Pred. No. 2.5e-49;
Matches 156; Conservative 42; Mismatches 100; Indels 19; Gaps 5;

QY 2 FVFPGGQPMGMRGRELIDSDVFRSVCCEAFAFYVDMVSVEQVLRDSDPAGLDRVD 61
DB 634 FVFPGGQGTQWAGMGRELIDSDSAVFAAAMAECEALSPYVDMSLDAVVRQAGATLERVD 693
QY 62 VVQPTLPAVWISLAALMRSGQVEPCAVLGHSLGEIAAAHVSGSLADAAVVTLSQA- 120
DB 694 VVQPTLPAVWISLAALMRSGQVEPCAVLGHSLGEIAAAHVSGSLADAAVVTLSKSI 753
QY 121 QTTLAGTGAIVSAATPDELLPRIAPWTEENPARLAVAANVGPSTVSGAREAVADIVA 180
DB 754 AAHLAGKGMISLALSEDAVLERLAGFD-----GLSVAANVGPATVSGDPVOIEELAR 808
QY 181 DLTAQVTRMIPVDVPAHSPLMYAIEERVVSGLLPTTPPSRIPFHSVTCGRDLTREL 240
DB 809 ACEADGVRARVITPDVASHRSQVEIISESLAEVLAGLSQAPRVFPFSTLEGAWITEPVL 868
QY 241 DAAVYRNMSSTVRFEPAR-LILQOGPKTVEVSPHVLTMGLQELAPDLGDTTGTADT 299
DB 869 DGGVYRNLRHVRGFAPAVETLATDGEFTHFVEVSAHPVLTMLP-----GTVTGLA-- 920
QY 300 VIMGTLRGGQTLDHFL 316
DB 921 ----TLRRDNGGQDRLV 933

RESULT 7

US-09-988-384B-31
; Sequence 31, Application US/09988384B
; Publication No. US20030073824M1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-31

Query Match 43.0%; Score 696.5; DB 9; Length 4613;
Best Local Similarity 49.2%; Pred. No. 2.5e-49;
Matches 156; Conservative 42; Mismatches 100; Indels 19; Gaps 5;

QY 2 FVFPGGQPMGMRGRELIDSDVFRSVCCEAFAFYVDMVSVEQVLRDSDPAGLDRVD 61
DB 634 FVFPGGQGTQWAGMGRELIDSDSAVFAAAMAECEALSPYVDMSLDAVVRQAGATLERVD 693
QY 62 VVQPTLPAVWISLAALMRSGQVEPCAVLGHSLGEIAAAHVSGSLADAAVVTLSQA- 120
DB 694 VVQPTLPAVWISLAALMRSGQVEPCAVLGHSLGEIAAAHVSGSLADAAVVTLSKSI 753

QY 121 QTTLAGTGAIVSAATPDELLPRIAPWTEENPARLAVAANVGPSTVSGAREAVADIVA 180
DB 754 AAHLAGKGMISLALSEDAVLERLAGFD-----GLSVAANVGPATVSGDPVOIEELAR 808
QY 181 DLTAQVTRMIPVDVPAHSPLMYAIEERVVSGLLPTTPPSRIPFHSVTCGRDLTREL 240
DB 809 ACEADGVRARVITPDVASHRSQVEIISESLAEVLAGLSQAPRVFPFSTLEGAWITEPVL 868
QY 241 DAAVYRNMSSTVRFEPAR-LILQOGPKTVEVSPHVLTMGLQELAPDLGDTTGTADT 299
DB 869 DGGVYRNLRHVRGFAPAVETLATDGEFTHFVEVSAHPVLTMLP-----GTVTGLA-- 920
QY 300 VIMGTLRGGQTLDHFL 316
DB 921 ----TLRRDNGGQDRLV 933

RESULT 8

US-09-836-821-31
; Sequence 31, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-836-821-31

Query Match 43.0%; Score 696.5; DB 9; Length 4613;
Best Local Similarity 49.2%; Pred. No. 2.5e-49;
Matches 156; Conservative 42; Mismatches 100; Indels 19; Gaps 5;

QY 2 FVFPGGQPMGMRGRELIDSDVFRSVCCEAFAFYVDMVSVEQVLRDSDPAGLDRVD 61
DB 634 FVFPGGQGTQWAGMGRELIDSDSAVFAAAMAECEALSPYVDMSLDAVVRQAGATLERVD 693
QY 62 VVQPTLPAVWISLAALMRSGQVEPCAVLGHSLGEIAAAHVSGSLADAAVVTLSQA- 120
DB 694 VVQPTLPAVWISLAALMRSGQVEPCAVLGHSLGEIAAAHVSGSLADAAVVTLSKSI 753
QY 121 QTTLAGTGAIVSAATPDELLPRIAPWTEENPARLAVAANVGPSTVSGAREAVADIVA 180
DB 754 AAHLAGKGMISLALSEDAVLERLAGFD-----GLSVAANVGPATVSGDPVOIEELAR 808
QY 181 DLTAQVTRMIPVDVPAHSPLMYAIEERVVSGLLPTTPPSRIPFHSVTCGRDLTREL 240
DB 809 ACEADGVRARVITPDVASHRSQVEIISESLAEVLAGLSQAPRVFPFSTLEGAWITEPVL 868
QY 241 DAAVYRNMSSTVRFEPAR-LILQOGPKTVEVSPHVLTMGLQELAPDLGDTTGTADT 299
DB 869 DGGVYRNLRHVRGFAPAVETLATDGEFTHFVEVSAHPVLTMLP-----GTVTGLA-- 920
QY 300 VIMGTLRGGQTLDHFL 316
DB 921 ----TLRRDNGGQDRLV 933

RESULT 9

US-09-861-289-31
; Sequence 31, Application US/09861289
; Patent No. US20020110897A1


```

; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-861-289-31.
```

```

Query Match      43.0%; Score 696.5; DB 10; Length 4613;
Best Local Similarity 49.2%; Pred. No. 2.5e-49;
Matches 156; Conservative 42; Mismatches 100; Indels 19; Gaps 5;
```

```

Cy 2 FVFPQGPQWPGMGRELIDASDVFRSVRACEAAPAPYVDMSEVOVLKDSPADGLDRVD 61
Db 634 FVFPQGPQWPGMGRELIDSSAVFAAAMACEALSPYVDMSEAVVROAPGAPTLERVD 693
Cy 62 VVQPTLFAVWISLAALMRSGVEPCAVLGHSLGEIAAAHVSGLSLDAARVVTLMSCA- 120
Db 694 VVQPTFAVWVSLARVWQHGVTPQAVVGHSGEIAAAYVAGLSLDDAARVVTLRKSI 753
Cy 121 QTTLAGTALVSAATPDELPRIPAWTEBNPARLAVAAVNGRSTVSGAREAVADLVA 180
Db 754 AAHLAGGGMLSLASEDAVLERLAGPD-----GLSVAAVNGRATVSSGDPVQIEELAR 808
Cy 181 DLTAQVTRTMIYVDVPAHSPLMYAIEERVVSGLLPTPRSPRIPFHSSVTGRLDTREL 240
Db 809 ACEADGVRARVIVDVASHRQVEIIESELAEVLAGISPOAPRVFPFSTLEGAWITEPVL 868
Cy 241 DAAVYRNNSSTVRFEPAR-LLLQGPKTVEVMSPPHVLTMGLQELAPDLGDTTGTADT 299
Db 869 DGGVWYRNLRHRVGFAPAVETLATDEGFTHFVEVSAHPVLTMLP-----GTVTGLA-- 920
Cy 300 VIMGTLRGGGTLDHFL 316
Db 921 -----TLRRDNGGQDRLV 933
```

```

RESULT 10
US-09-860-846-6
; Sequence 6, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 11877
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-6
```

```

Query Match      43.0%; Score 696.5; DB 9; Length 11877;
Best Local Similarity 49.2%; Pred. No. 8.8e-49;
```

```

Matches 156; Conservative 42; Mismatches 100; Indels 19; Gaps 5;
Cy 2 FVFPQGPQWPGMGRELIDSDVFRSVRACEAAPAPYVDMSEVOVLKDSPADGLDRVD 61
Db 970 FVFPQGPQWPGMGRELIDSSAVFAAAMACEALSPYVDMSEAVVROAPGAPTLERVD 1029
Cy 62 VVQPTLFAVWISLAALMRSGVEPCAVLGHSLGEIAAAHVSGLSLDAARVVTLMSCA- 120
Db 1030 VVQPTFAVWVSLARVWQHGVTPQAVVGHSGEIAAAYVAGLSLDDAARVVTLRKSI 1089
Cy 121 QTTLAGTALVSAATPDELPRIPAWTEBNPARLAVAAVNGRSTVSGAREAVADLVA 180
Db 1090 AAHLAGGGMLSLASEDAVLERLAGPD-----GLSVAAVNGRATVSSGDPVQIEELAR 1144
Cy 181 DLTAQVTRTMIYVDVPAHSPLMYAIEERVVSGLLPTPRSPRIPFHSSVTGRLDTREL 240
Db 1145 ACEADGVRARVIVDVASHRQVEIIESELAEVLAGISPOAPRVFPFSTLEGAWITEPVL 1204
Cy 241 DAAVYRNNSSTVRFEPAR-LLLQGPKTVEVMSPPHVLTMGLQELAPDLGDTTGTADT 299
Db 1205 DGGVWYRNLRHRVGFAPAVETLATDEGFTHFVEVSAHPVLTMLP-----GTVTGLA-- 1256
Cy 300 VIMGTLRGGGTLDHFL 316
Db 1257 -----TLRRDNGGQDRLV 1269
```

```

RESULT 11
US-09-836-821-6
; Sequence 6, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 11877
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-836-821-6
```

```

Query Match      43.0%; Score 696.5; DB 9; Length 11877;
Best Local Similarity 49.2%; Pred. No. 8.8e-49;
Matches 156; Conservative 42; Mismatches 100; Indels 19; Gaps 5;
Cy 2 FVFPQGPQWPGMGRELIDSDVFRSVRACEAAPAPYVDMSEVOVLKDSPADGLDRVD 61
Db 970 FVFPQGPQWPGMGRELIDSSAVFAAAMACEALSPYVDMSEAVVROAPGAPTLERVD 1029
Cy 62 VVQPTLFAVWISLAALMRSGVEPCAVLGHSLGEIAAAHVSGLSLDAARVVTLMSCA- 120
Db 1030 VVQPTFAVWVSLARVWQHGVTPQAVVGHSGEIAAAYVAGLSLDDAARVVTLRKSI 1089
Cy 121 QTTLAGTALVSAATPDELPRIPAWTEBNPARLAVAAVNGRSTVSGAREAVADLVA 180
Db 1090 AAHLAGGGMLSLASEDAVLERLAGPD-----GLSVAAVNGRATVSSGDPVQIEELAR 1144
Cy 181 DLTAQVTRTMIYVDVPAHSPLMYAIEERVVSGLLPTPRSPRIPFHSSVTGRLDTREL 240
Db 1145 ACEADGVRARVIVDVASHRQVEIIESELAEVLAGISPOAPRVFPFSTLEGAWITEPVL 1204
Cy 241 DAAVYRNNSSTVRFEPAR-LLLQGPKTVEVMSPPHVLTMGLQELAPDLGDTTGTADT 299
Db 1205 DGGVWYRNLRHRVGFAPAVETLATDEGFTHFVEVSAHPVLTMLP-----GTVTGLA-- 1256
```

QY 300 VIMGTLRRGGCTLDHFL 316
1257 -----TLRRDNGGQDRLV 1269

RESULT 12

US-09-861-289-6
Sequence 6, Application US/09861289
Patent No. US20020110897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/861,289
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-861-289-6

Query Match 43.0%; Score 696.5; DB 10; Length 11877;
Best Local Similarity 49.2%; Pred. No. 8.8e-49;
Matches 156; Conservative 42; Mismatches 100; Indels 19; Gaps 5;

QY 2 FVFPGGQPGWPGMGRELLDASDVRESVRACEAFAFAYVDMVSEVOYL RDSPPDAGLDRLVD 61
970 FVFPGGQPGWPGMGRELLDSSAVFAAABCEALSPYDMSLEAVVRQAPGAPTLERVD 1029
DB 62 VVQPTLFVAVMISLAALMRSGVEPCAVLGHSIGETIAAAHVSGSLADAAVVTLMSQA- 120
1030 VVQPTLFVAVMISLARVWQHGHGVTPOAVVGHSGEIAAAVAGALSDDAAVVTLRSKSI 1089
QY 121 QTTLAGTALVSAATPDELLPRIAPWTEENPARIAVAANGPRSTVSGAREAVDLVA 180
1090 AAHLAGKGMLSLALSEDAVLERLAGFD-----GLSVAANGPTATVSGDPVQIEELAR 1144
DB 181 DLTAQVTRMIPVDVPAHSPLMAIERVVSGLPTPRSRIPFHSVTCGLDTRREL 240
1145 ACEADGVARAVIPVDVASHRQVEITISELAELVLAGSPQAPRVFPFSTLEGAMITEPVL 1204
QY 241 DAAVYRNMSSTVFEFPAAR-LILQOGPKTFVEMSPHVLTMGLQELAPDLGDTTGTADT 299
DB 1205 DGGWYNNLRHRVGFAPAVETLATDGGTFHVEVSAPVLTMLP-----GTVTGIA-- 1256
QY 300 VIMGTLRRGGCTLDHFL 316
DB 1257 -----TLRRDNGGQDRLV 1269

RESULT 13

US-09-988-384B-6
Sequence 6, Application US/09988384B
Publication No. US2003007824A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.536US1
CURRENT APPLICATION NUMBER: US/09/988,384B
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: PCT/US99/14398
PRIOR FILING DATE: 1999-06-25

PRIOR APPLICATION NUMBER: US 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 53
SEQ ID NO 6
LENGTH: 12199
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-988-384B-6

Query Match 43.0%; Score 696.5; DB 9; Length 12199;
Best Local Similarity 49.2%; Pred. No. 9.1e-49;
Matches 156; Conservative 42; Mismatches 100; Indels 19; Gaps 5;

QY 2 FVFPGGQPGWPGMGRELLDASDVRESVRACEAFAFAYVDMVSEVOYL RDSPPDAGLDRLVD 61
DB 1292 FVFPGGQPGWPGMGRELLDSSAVFAAABCEALSPYDMSLEAVVRQAPGAPTLERVD 1351
QY 62 VVQPTLFVAVMISLAALMRSGVEPCAVLGHSIGETIAAAHVSGSLADAAVVTLMSQA- 120
DB 1352 VVQPTLFVAVMISLARVWQHGHGVTPOAVVGHSGEIAAAVAGALSDDAAVVTLRSKSI 1411
QY 121 QTTLAGTALVSAATPDELLPRIAPWTEENPARIAVAANGPRSTVSGAREAVDLVA 180
DB 1412 AAHLAGKGMLSLALSEDAVLERLAGFD-----GLSVAANGPTATVSGDPVQIEELAR 1466
QY 181 DLTAQVTRMIPVDVPAHSPLMAIERVVSGLPTPRSRIPFHSVTCGLDTRREL 240
DB 1467 ACEADGVARAVIPVDVASHRQVEITISELAELVLAGSPQAPRVFPFSTLEGAMITEPVL 1526
QY 241 DAAVYRNMSSTVFEFPAAR-LILQOGPKTFVEMSPHVLTMGLQELAPDLGDTTGTADT 299
DB 1527 DGGWYNNLRHRVGFAPAVETLATDGGTFHVEVSAPVLTMLP-----GTVTGIA-- 1578
QY 300 VIMGTLRRGGCTLDHFL 316
DB 1579 -----TLRRDNGGQDRLV 1591

RESULT 14

US-09-860-846-33
Sequence 33, Application US/09860846
Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/860,846
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-860-846-33

Query Match 42.1%; Score 681.5; DB 9; Length 3739;
Best Local Similarity 48.9%; Pred. No. 3.5e-48;
Matches 152; Conservative 42; Mismatches 98; Indels 19; Gaps 5;

QY 2 FVFPGGQPGWPGMGRELLDASDVRESVRACEAFAFAYVDMVSEVOYL RDSPPDAGLDRLVD 61
DB 2069 FVFPGGQPGWPGMGRELLDSSAVFAAABCEALSPYDMSLEAVVRQAPGAPTLERVD 2128
QY 62 VVQPTLFVAVMISLAALMRSGVEPCAVLGHSIGETIAAAHVSGSLADAAVVTLMSQA- 120
DB 2129 VVQPTLFVAVMISLARVWQHGHGVTPOAVVGHSGEIAAAVAGALSDDAAVVTLRSKSI 2188

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:54:32 / Search time 5.74347 Seconds
(without alignments)
1618.821 Million cell updates/sec

Title: US-09-914-286-3_COPY_29_344

Perfect score: 1618

Sequence: 1 VFVFPQGGPQWPGMGRRLD.....ADTVIMGTLRGQGLDHL 316

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: Issued_Patents_AA:*
2: /cgn2_6/p/ptodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/p/ptodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/p/ptodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/p/ptodata/1/1aa/PTUS.COMB.pep:*
6: /cgn2_6/p/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	819	50.6	344	US-08-858-003-33	Sequence 33, Appl
2	819	50.6	344	US-09-078-166-33	Sequence 33, Appl
3	819	50.6	344	US-08-997-467-33	Sequence 33, Appl
4	809.5	50.0	3724	US-08-804-227C-10	Sequence 10, Appl
5	809.5	50.0	3724	US-08-804-198-4	Sequence 4, Appl
6	776.5	48.0	3567	US-07-642-734C-4	Sequence 4, Appl
7	776.5	48.0	3567	US-08-439-009A-4	Sequence 4, Appl
8	775.5	47.9	829	US-09-413-814-105	Sequence 105, App
9	765	47.3	2595	US-09-036-987A-2	Sequence 2, Appl
10	765	47.3	2595	US-09-370-700-2	Sequence 2, Appl
11	762	47.1	3729	US-08-804-227C-6	Sequence 6, Appl
12	750	46.4	5588	US-09-036-987A-6	Sequence 6, Appl
13	750	46.4	5588	US-09-370-700-6	Sequence 6, Appl
14	726	44.9	1580	US-08-804-227C-11	Sequence 11, Appl
15	726	44.9	1580	US-08-804-198-5	Sequence 5, Appl
16	722	44.6	3491	US-07-642-734C-2	Sequence 2, Appl
17	722	44.6	3491	US-08-439-009A-2	Sequence 2, Appl
18	714.5	44.2	3816	US-09-428-517-3	Sequence 3, Appl
19	712	44.0	4150	US-09-428-517-2	Sequence 2, Appl
20	707	43.7	3519	US-09-428-517-4	Sequence 4, Appl
21	702.5	43.4	3170	US-07-642-734C-5	Sequence 5, Appl
22	702.5	43.4	3170	US-08-439-009A-5	Sequence 5, Appl
23	696.5	43.0	4551	US-09-320-878-1	Sequence 1, Appl
24	696.5	43.0	4613	US-09-105-537-31	Sequence 31, Appl
25	686.5	43.0	11877	US-09-105-537-6	Sequence 6, Appl
26	681.5	42.1	3739	US-09-320-878-2	Sequence 2, Appl
27	681.5	42.1	3739	US-09-105-537-33	Sequence 33, Appl

28	674.5	41.7	3170	US-09-036-987A-4	Sequence 4, Appl
29	674.5	41.7	3170	US-09-370-700-4	Sequence 4, Appl
30	674	41.7	1562	US-09-320-878-3	Sequence 3, Appl
31	674	41.7	1562	US-09-105-537-35	Sequence 35, Appl
32	669.5	41.4	6095	US-09-144-085-2	Sequence 2, Appl
33	667.5	41.3	1346	US-09-320-878-4	Sequence 4, Appl
34	667.5	41.3	1346	US-09-105-537-34	Sequence 34, Appl
35	667	41.2	345	US-08-858-003-34	Sequence 34, Appl
36	667	41.2	345	US-08-858-003-34	Sequence 34, Appl
37	667	41.2	345	US-08-997-467-34	Sequence 34, Appl
38	662	40.9	5087	US-09-144-085-1	Sequence 1, Appl
39	652.5	40.3	841	US-09-413-814-107	Sequence 107, App
40	627.5	38.8	2890	US-09-413-814-67	Sequence 67, Appl
41	627.5	38.8	3798	US-09-335-409-6	Sequence 6, Appl
42	627.5	38.8	3798	US-09-568-102-6	Sequence 6, Appl
43	627.5	38.8	3798	US-09-567-969-6	Sequence 6, Appl
44	627.5	38.8	3798	US-09-568-480-6	Sequence 6, Appl
45	627.5	38.8	3798	US-09-568-486-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-858-003-33
Sequence 33, Application US/08858003
Patent No. 6060234
GENERAL INFORMATION:
APPLICANT: Katz, Leonard
APPLICANT: Stasel, Diane L.
APPLICANT: Summers Jr., Richard G.
APPLICANT: Ruan, Xiaocan
APPLICANT: Pereda-Lopez, Ana
APPLICANT: Kakavas, Stephan J.
TITLE OF INVENTION: NOVEL POLYPEPTIDE DERIVATIVES
TITLE OF INVENTION: AND RECOMBINANT METHODS FOR MAKING SAME
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Rd.
CITY: Abbott Park
STATE: Illinois
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/858,003
FILING DATE: 16-MAY-1979
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Diane Casuto
REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 4952.US.P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (847)-938-3137
TELEFAX: (847)-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6060234e
US-08-858-003-33

Query Match 50.6%; Score 819; DB 3; Length 344;
Best Local Similarity 54.0%; Pred. No. 4.9e-73;
Matches 168; Conservative 38; Mismatches 95; Indels 10; Gaps 3;

QY 1 VFVPPGGQSPWPMGKRELLDSDVFRRESVACGAAPFYVWVSVEQYLRSDPAPGLDRV 60

DB 5 VFVPPGGQSQWAGAEGLERSGAFRSADSCDAALRPYLGWSVLSVLRGEPDAPSLDRV 64

QY 61 DVVQPTLFAMWISLAALMRSGVEPCAVLGSHGELIAAAHVSGGLSLADAAVYVTLMSQA 120

DB 65 DVVQPVLFMTWVSLAAVWRALGVEPAVAVHSGGEIAAAHVAGLSLSDSARIVALSRA 124

QY 121 QTTIAGTALVSAATPDELLPRIAPWTEEDNPARLAAVAVNGPRSTVSGAREAVDLVA 180

DB 125 WLGAGKGMVAVPMPEELRPRLVTWGD---RLAVAAVNSPGSCAVAGDPALAEIVA 180

QY 181 DLTAQVTRMIP-VDVPAHSPLMYAIEERVVSGLPITPRBSRIPPHSSVTGGRDLTRE 239

DB 181 LITGEGVHARPIPVDTAGHSPOVDALRAHLEVLAPVAPRPADIPIFYSTVTGGLDGT 240

QY 240 LDDAAYWRNNSSTVREFPARLLLOQPKTFVMSPPHVLTMGLQELAPDLGDTGTADT 299

DB 241 LDATYWRNMRKEVEFERATRALIADGHVFLTSPHMLAVALEQVTTDAG-----TDA 295

QY 300 VINGTLRRGQG 310

DB 296 AVLGTLRRRHG 306

RESULT 2

US-09-078-166-33

Sequence 33, Application US/09078166

Patent No. 6063561

GENERAL INFORMATION:

APPLICANT: Katz, Leonard

APPLICANT: Stassl, Diane L.

APPLICANT: Summers Jr., Richard G.

APPLICANT: Ruan, Xiaocan

APPLICANT: Pereda-Lopez, Ana

APPLICANT: Kakavas, Stephan J.

TITLE OF INVENTION: NOVEL POLYKETIDE DERIVATIVES

TITLE OF INVENTION: AND RECOMBINANT METHODS FOR MAKING SAME

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Rd.

CITY: Abbott Park

STATE: Illinois

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/078,166

FILING DATE: 16-MAY-1979

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Diane Casuto

REGISTRATION NUMBER: P-40,943

REFERENCE/DOCKET NUMBER: 4952, US, P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (847)-938-3137

TELEFAX: (847)-938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 344 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6063561e

US-09-078-166-33

Query Match 50.6%; Score 819; DB 3; Length 344;

Best Local Similarity 54.0%; Pred. No. 4.9e-73;

Matches 168; Conservative 38; Mismatches 95; Indels 10; Gaps 3;

QY 1 VFVPPGGQSPWPMGKRELLDSDVFRRESVACGAAPFYVWVSVEQYLRSDPAPGLDRV 60

DB 5 VFVPPGGQSQWAGAEGLERSGAFRSADSCDAALRPYLGWSVLSVLRGEPDAPSLDRV 64

QY 61 DVVQPTLFAMWISLAALMRSGVEPCAVLGSHGELIAAAHVSGGLSLADAAVYVTLMSQA 120

DB 65 DVVQPVLFMTWVSLAAVWRALGVEPAVAVHSGGEIAAAHVAGLSLSDSARIVALSRA 124

QY 121 QTTIAGTALVSAATPDELLPRIAPWTEEDNPARLAAVAVNGPRSTVSGAREAVDLVA 180

DB 125 WLGAGKGMVAVPMPEELRPRLVTWGD---RLAVAAVNSPGSCAVAGDPALAEIVA 180

QY 181 DLTAQVTRMIP-VDVPAHSPLMYAIEERVVSGLPITPRBSRIPPHSSVTGGRDLTRE 239

DB 181 LITGEGVHARPIPVDTAGHSPOVDALRAHLEVLAPVAPRPADIPIFYSTVTGGLDGT 240

QY 240 LDDAAYWRNNSSTVREFPARLLLOQPKTFVMSPPHVLTMGLQELAPDLGDTGTADT 299

DB 241 LDATYWRNMRKEVEFERATRALIADGHVFLTSPHMLAVALEQVTTDAG-----TDA 295

QY 300 VINGTLRRGQG 310

DB 296 AVLGTLRRRHG 306

RESULT 3

US-08-997-467-33

Sequence 33, Application US/08997467

Patent No. 6200813

GENERAL INFORMATION:

APPLICANT: Katz, Leonard

APPLICANT: Stassl, Diane L.

APPLICANT: Summers Jr., Richard G.

APPLICANT: Ruan, Xiaocan

APPLICANT: Pereda-Lopez, Ana

APPLICANT: Kakavas, Stephan J.

TITLE OF INVENTION: NOVEL POLYKETIDE DERIVATIVES

TITLE OF INVENTION: AND RECOMBINANT METHODS FOR MAKING SAME

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Rd.

CITY: Abbott Park

STATE: Illinois

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/997,467

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/858,003

FILING DATE: 16-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Diane Casuto

REGISTRATION NUMBER: P-40,943

REFERENCE/DOCKET NUMBER: 4952, US, P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (847)-938-3137
 TELEFAX: (847)-938-2623
 TELEX:
 INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 344 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: No. 6200813e
 US-08-997-467-33

Query Match 50.6%; Score 819; DB 4; Length 344;
 Best Local Similarity 54.0%; Pred. No. 4.9e-73;
 Matches 168; Conservative 38; Mismatches 95; Indels 10; Gaps 3;

QY 1 VFVPPGCGPQPMGREL--DASDFRESVACAAAPYVDMSVEQYLRSPPDAPGLDRV 60
 DB 5 VFVPPGCGSQWAGAEGLERSGAFRBAADSCDALPPLGWSVSLVLRGEPDAPSLDRV 64
 QY 61 DVVQPTLFAVMISLAALMRSQVPCAVLGHSLGEIAAAHVSGGLSLADAARVYTLWSQA 120
 DB 65 DVVQPVPLTMMVSLAAVWRALGVPAVVGHSQGEIAAAHVAGLSLDDSRIVALSRA 124
 QY 121 QTTLAGTALVSVAATPDELLPRIAPWTEBNPARLAAVAVNGPRSTVSGAREAVADLV 180
 DB 125 WLGLAGKGVAVVPAEELRLVTWGD---RLAVAAVNSPSCAVAGDPBALALVA 180
 QY 181 DLRLAQRTMIP-VDVPAHSPLMYAIEERVVSGLLPTTRPSRIIPHSSTVGGRLDTR 239
 DB 181 LTLGEGHARPIPGVDTAGHSPQVDALRAHLBVLAVAPAPADIPYSTVTGGLDGT 240
 QY 240 LDAAYVRNNSSTVRPEPARLLLOQPKTFVEMSPHVLTMGLQELAPLDGTTGTADT 299
 DB 241 LDATYTRNRKEPVEFRATRALIADGHDFLETSPPMLAVALEQVTTAG-----TDA 295
 QY 300 VINGTLRRGQ 310
 DB 296 AVLGTLRRRHG 306

RESULT 4
 US-08-804-227C-10
 Sequence 10, Application US/08804227C
 Patent No. 5876991
 GENERAL INFORMATION:
 APPLICANT: Dehoff, Bradley S.
 APPLICANT: Kubloos, Stuart A.
 APPLICANT: Roosteck, Paul R., Jr.
 APPLICANT: Sutton, Kimberly L.
 TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: THOMAS G. PLANT 1501
 STREET: LILLY CORPORATE CENTER
 CITY: INDIANAPOLIS
 STATE: IN
 COUNTRY: USA
 ZIP: 46285
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: ASCII(DOS) Text only
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/804,227C
 FILING DATE: February 21, 1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Plant, Thomas G.
 REGISTRATION NUMBER: 35,784
 REFERENCE/DOCKET NUMBER: X-8221
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-2459
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3724 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-804-227C-10

Query Match 50.0%; Score 809.5; DB 2; Length 3724;
 Best Local Similarity 54.1%; Pred. No. 1.5e-70;
 Matches 170; Conservative 37; Mismatches 94; Indels 13; Gaps 4;

QY 1 VFVPPGCGPQPMGREL---DASDFRESVACAAAPYVDMSVEQYLRSPPDAPGL 57
 DB 2089 VFVPPGCGSQWPMAGGLARSSGSGSFLETAACDLARPHLGWSVLDVLRGEPGSL 2148
 QY 58 DRVDVQPTLFAVMISLAALMRSQVPCAVLGHSLGEIAAAHVSGGLSLADAARVYTLW 117
 DB 2149 DRVDVQPVPLTMMVSLAETWRSIGVPAVVGHSQGEIAAAHVAGALTDDAARIVALLR 2208
 QY 118 SQHQTLAGTALVSVAATPDELLPRIAPWTEBNPARLAAVAVNGPRSTVSGAREAVAD 177
 DB 2209 SQAMRLAGKGVAVVTLSEDRPLRLPWS---RLAVAAVNGPCTCAVSDPDALAE 2264
 QY 178 LVADLTAAQVTRMIP-VDVPAHSPLMYAIEERVVSGLLPTTRPSRIIPHSSTVGGRLD 236
 DB 2265 LVALGEGHARPIPGVDTAGHSPQVDTEALRLKVLAVAPARTDIPYSTVTGGLID 2324
 QY 237 TRELDAAYVRNNSSTVRPEPARLLLOQPKTFVEMSPHVLTMGLQELAPLDGTTGT 296
 DB 2325 TAEILDADYWRNRKEPVEFRATRALIADGHDFLESSPHMLAVSLQETISDAG----- 2379
 QY 297 ADVINGTLRRGQ 310
 DB 2380 SPAVLDGLRRGQ 2393

RESULT 5
 US-08-804-198-4
 Sequence 4, Application US/08804198
 Patent No. 5945320
 GENERAL INFORMATION:
 APPLICANT: Burgess, Stanley G.
 APPLICANT: Kubloos, Stuart A.
 APPLICANT: Rao, Nagaraja R.
 APPLICANT: Richardson, Mark A.
 APPLICANT: Roosteck, Paul R., Jr.
 TITLE OF INVENTION: PLATENOILIDE SYNTHASE GENE
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: PAUL R. CANTRELL 1138
 STREET: LILLY CORPORATE CENTER
 CITY: INDIANAPOLIS
 STATE: IN
 COUNTRY: USA
 ZIP: 46285
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Macintosh
 OPERATING SYSTEM: Macintosh 7.0
 SOFTWARE: Microsoft Word 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/804,198
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: CANTRELL, PAUL R.
 REGISTRATION NUMBER: 36,470
 REFERENCE/DOCKET NUMBER: P9113
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 317-276-3885
 INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 3724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-4

Query Match 50.0%; Score 809.5; DB 2; Length 3724;
Best Local Similarity 54.1%; Pred. No. 1.5e-70;
Matches 170; Conservative 37; Mismatches 94; Indels 13; Gaps 4;

QY 1 VFVPPGPGPMPGKRELL--DASDFRESVRACEAFAFYDMSVEQVLRSDPADGL 57
DB 2089 VFVPPGPGSQMPENADGILARSSSGSEFLFARACDLARPLGMSVLDVLRREPGAPSL 2148
QY 58 DRVVVQPTLFAVWISLAALMRSGVEPCAVLGHSIGETIAAHSVSGSLADARVVTLM 117
DB 2149 DRVVVQPTLFTWVWSLAETWRSIGVEPAVAVGSHSGEIAAAYVAGALTDDAARVALR 2208
QY 118 SQAQTLAGTALVSAATPDELLPRIAPWTEEDNPRLAANAANGPRSTVSGAREAVD 177
DB 2209 SQALRLRAGKGMVAVTSEDDLRLRLEPMSD---RLAANAANGPETCAVSGCPDLAE 2264
QY 178 LVADLTAAQVTRMIP-VDVPAHSPLMYAIEERVVSGLLPTPPPSRIPFHSVVTGRLD 236
DB 2265 LVAEIGAEVHARPIPGVDTAGHSPOVDTLBAHLKVLAPVAPRSDIPFYSTVTGGLID 2324
QY 237 TREIDAAVYWNMSSTVFEPAARLLLOQGPKEVMSPPHVLTMGLQELAPDLGDTTGT 296
DB 2325 TAEIDADVYWNMBEPFEQATRALIADGHVFLSSPHMLAVSLQETISDXG----- 2379
QY 297 ADTVIMGTLRRGQG 310
DB 2380 SPAVLGTLRRGQG 2393

RESULT 6
US-07-642-734C-4
Sequence 4, Application US/07642734C
Patent No. 5824513
GENERAL INFORMATION:
APPLICANT: Katz, L
APPLICANT: Donadio, S
APPLICANT: Mcalpine, J B
TITLE OF INVENTION: Recombinant DNA Method for Producing
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman
STREET: Abbott Laboratories D377/APeD-2 One Abbott
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/642,734C
FILING DATE: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Danckerts, Andreas M
REGISTRATION NUMBER: 32652
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-4

LENGTH: 3567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-642-734C-4

Query Match 48.0%; Score 776.5; DB 2; Length 3567;
Best Local Similarity 50.0%; Pred. No. 2.7e-67;
Matches 158; Conservative 44; Mismatches 107; Indels 7; Gaps 2;

QY 1 VFVPPGPGPMPGKRELLDASDFRESVRACEAFAFYDMSVEQVLRSDPADGLDRV 60
DB 560 VFLEPPGPGSQMAGGELLSSPVFAKTRACDSNAPMDWVSDVLRQAPGGLDRV 619
QY 61 DVVQPTLFAVWISLAALMRSGVEPCAVLGHSIGETIAAHSVSGSLADARVVTLM 120
DB 620 DVVQPTLFAVWISLAELMRSGVEPAVAVGSHSGEIAAAYVAGALTIEDPAKLVGSR 679
QY 121 QTTLAGTALVSAATPDELLPRIAPWTEEDNPRLAANAANGPRSTVSGAREAVD 180
DB 680 MRSLSGSGMAAVALGSAARERLRPMQD---RLSVAANAANGPRSVVSGEPGLARFSE 735
QY 181 DLTPAQVTRMIPVDVPAHSPLMYAIEERVVSGLLPTPPPSRIPFHSVVTGRLDREL 240
DB 736 DCAAGIRVADIDVDYASHSPQIERVEEELTTGDIAPRPARVTFESRSRSMGTET 795
QY 241 DAAVYWNMSSTVFEPAARLLLOQGPKEVMSPPHVLTMGLQELAPDLGDTTGTADTV 300
DB 796 DAAVYWNMSSTVFEPAARLLLOQGPKEVMSPPHVLTMGLQELAPDLGDTTGTADTV 300
QY 301 IMGTLRRGQTLDFL 316
DB 853 VVGSILHRDGDLSAFL 868

RESULT 7
US-08-439-009A-4
Sequence 4, Application US/08439009A
Patent No. 6004787
GENERAL INFORMATION:
APPLICANT: Katz, L
APPLICANT: Donadio, S
APPLICANT: Mcalpine, J B
TITLE OF INVENTION: Method of Directing Biosynthesis of
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven F. Weinstock
STREET: Abbott Laboratories D377/APeD-2 One Abbott
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,009A
FILING DATE: 11-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Diane
REGISTRATION NUMBER: 40,943
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3567 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-439-009A-4

Query Match 48.0%; Score 776.5; DB 3; Length 3567;
Best Local Similarity 50.0%; Pred. No. 2.7e-67;
Matches 158; Conservative 44; Mismatches 107; Indels 7; Gaps 2;

QY 1 VFVPPGQGPQWPGMGBELLDDASDVFRSVCBAAPAPYVMSVEOVLKDSPPAPGLDRV 60
DB 560 VFLEPPGQSGQWAGAGAILSSPFVAGKIRACDSMAFMQMKISDVLKRAFGAPGLDRV 619
QY 61 DVVQPTLFVAVMISLAALMRSGVPCAVLGHSLGEIAAAHVSGGLSLADAAVYVTLWSQA 120
DB 620 DVVQPTLFVAVMISLAALMRSGVPCAVLGHSLGEIAAAHVSGGLSLADAAVYVTLWSQA 679
QY 121 QTTIAGTALVSVAATPEDELLPRIAPWTEPNPALAAVAVNGPSTVSGAREVADLVA 180
DB 680 MRSLSGEGGMAAVALGEAAVVERLRPMOD---RLSVAAVNGPSTVSGAREVADLVA 735
QY 181 DLTAQVTRMIPVDVPAHSPMLVAIEERVVSGLLPTPRPSRIPFHSSVTGRLDTREL 240
DB 736 DCAABEGIRARDIDVDYSHSPQIRVREBELLETTGDIAPRPARTFSTVESRMDTREL 795
QY 241 DAAVYRNMSSTVREPPAARLLLOQPKTFVMSPPHVLTMGLQELAPDLDTGTADTV 300
DB 796 DAAVYRNLRRTVAFDAVTRLAESGYDAFIEVSPHVVQAVEA---VEADGADAV 852
QY 301 IMGTLRGQGTLDHFL 316
DB 853 VVSLHRDGGDLAFL 868

RESULT 8

US-09-413-814-105
Sequence 105; Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Geneslaucht fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Bayer, Stefan
APPLICANT: Bloeker, Helmut
APPLICANT: Cino, Paul M.
APPLICANT: Dougherty, Brian A.
APPLICANT: Goldberg, Steven L.
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 105
LENGTH: 829
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-105

Query Match 47.9%; Score 775.5; DB 4; Length 829;
Best Local Similarity 52.8%; Pred. No. 3.7e-68;
Matches 168; Conservative 46; Mismatches 91; Indels 13; Gaps 4;

QY 2 FVFPQGGQGPQWPGMGBELLDDASDVFRSVCBAAPAPYVMSVEOVLKDSPPAPGLDRV 59
DB 338 FVFPQGGQGPQWPGMGBELLDDASDVFRSVCBAAPAPYVMSVEOVLKDSPPAPGLDRV 397
QY 60 DVVQPTLFVAVMISLAALMRSGVPCAVLGHSLGEIAAAHVSGGLSLADAAVYVTLWSQ 119

DB 398 DVVQPTLFVAVMISLAALMRSGVPCAVLGHSLGEIAAAHVSGGLSLADAAVYVTLWSQ 457
QY 120 AOTIAGTALVSVAATPEDELLPRIAPWTEPNPALAAVAVNGPSTVSGAREVADLVA 179
DB 458 AITKLGRGMSAVELTAELELAPLGS---RLAIATAINSPPHALVSGDDPGADLVL 513
QY 180 ADLTAQVTRMIPVDVPAHSPMLVAIEERVVSGLLPTPRPSRIPFHSSVTGRLDTREL 239
DB 514 AELSGAQLFARKRVREYASHSAHVAIEERTMLEALOSIARPAVPLYSAVTGERLDGEA 573
QY 240 LDAVYRNMSSTVREPPAARLLLOQPKTFVMSPPHVLTMGLQELAPDL-GDTGTAD 298
DB 574 LGAHWYRNLRHTVREFAHATRALDDGHRFEVSPHVLTVALDETSLSSIDGAT---- 629
QY 299 TVIMGTLRGQGTLDHFL 316
DB 630 -VVSLRDEGDLERFL 645

RESULT 9

US-09-036-987A-2
Sequence 2; Application US/09036987A
Patent No. 6143526
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: BioSynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow AgroSciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-2

Query Match 47.3%; Score 765; DB 4; Length 2595;
Best Local Similarity 50.6%; Pred. No. 2.3e-66;
Matches 157; Conservative 51; Mismatches 92; Indels 10; Gaps 3;

QY 1 VFVPPGQGPQWPGMGBELLDDASDVFRSVCBAAPAPYVMSVEOVLKDSPPAPGLDRV 60

Db 527 VFVPPGSGQWIGMGRQLLETSEVFAAASMDCADALAPHLDMSLDVLRNAAGAHLDHD 586
 QY 61 DVVQPTLFAMWISLAALMRSGVPCAVLGHSLGEIAAAHVSGLSLADARVVTLSQA 120
 Db 587 DVVQPALFALWISLAELMRSGVPRVAVVGHSGEIAAACVAGALSVDAARVAVVSRLL 646
 QY 121 QTTLAGTALVSATPDELPRIPAPWTEENPARLAAVAVNGPRSTVSGAREAVADIVA 180
 Db 647 LITAGSGAMASLGHRAEVRQILLPMRD----RIGVAGVNGPSTVLVSGDREMAIELLA 702
 QY 181 DLTAQVTRTMIIPDVPAHSPMLVAIERVVSGLPIPTPRSRIPFHSVVGRLDTRRL 240
 Db 703 ECADREIRMRRIPEYASHSPHIEVRDELGLLAPVPRGSIPIVSTTGGDLID-RPM 761
 QY 241 DAAYWRMSSTVRFEPARALLLOOGPKTFVMSPHVLTWGLQELAPDLGDTTGTADTV 300
 Db 762 DADWYRNLRQPVLFEEAAVEALLKRGYDAFLEISPHVLTANTIOETAVRAG-----REV 816
 QY 301 IMGTLLRRGQG 310
 Db 817 ALGTLRRGEG 826

RESULT 10

US-09-370-700-2
 ; Sequence 2, Application US/09370700
 ; Patent No. 6274350
 ; GENERAL INFORMATION:
 ; APPLICANT: Balz, Richard H
 ; APPLICANT: Broughton, Mary C
 ; APPLICANT: Crawford, Kathryn P
 ; APPLICANT: Madhuri, Krishnamurthy
 ; APPLICANT: Turner, Jan R
 ; APPLICANT: Treadway, Patti J
 ; APPLICANT: Waldron, Clive
 ; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
 ; FILE REFERENCE: 50489 DIV1
 ; CURRENT APPLICATION NUMBER: US/09/370, 700
 ; CURRENT FILING DATE: 1999-08-09
 ; EARLIER APPLICATION NUMBER: US 09/36987
 ; EARLIER FILING DATE: 1998-03-09
 ; NUMBER OF SEQ ID NOS: 39
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 2595
 ; TYPE: PRT
 ; ORGANISM: Saccharopolyspora spinosa
 ; US-09-370-700-2

Query Match 47.3%; Score 765; DB 4; Length 2595;
 Best Local Similarity 50.6%; Pred. No. 2.3e-66;
 Matches 157; Conservative 51; Mismatches 92; Indels 10; Gaps 3;

QY 1 VFVPPGSGQWIGMGRQLLETSEVFAAASMDCADALAPHLDMSLDVLRNAAGAHLDHD 60
 Db 527 VFVPPGSGQWIGMGRQLLETSEVFAAASMDCADALAPHLDMSLDVLRNAAGAHLDHD 586
 QY 61 DVVQPTLFAMWISLAALMRSGVPCAVLGHSLGEIAAAHVSGLSLADARVVTLSQA 120
 Db 587 DVVQPALFALWISLAELMRSGVPRVAVVGHSGEIAAACVAGALSVDAARVAVVSRLL 646
 QY 121 QTTLAGTALVSATPDELPRIPAPWTEENPARLAAVAVNGPRSTVSGAREAVADIVA 180
 Db 647 LITAGSGAMASLGHRAEVRQILLPMRD----RIGVAGVNGPSTVLVSGDREMAIELLA 702
 QY 181 DLTAQVTRTMIIPDVPAHSPMLVAIERVVSGLPIPTPRSRIPFHSVVGRLDTRRL 240
 Db 703 ECADREIRMRRIPEYASHSPHIEVRDELGLLAPVPRGSIPIVSTTGGDLID-RPM 761
 QY 241 DAAYWRMSSTVRFEPARALLLOOGPKTFVMSPHVLTWGLQELAPDLGDTTGTADTV 300
 Db 762 DADWYRNLRQPVLFEEAAVEALLKRGYDAFLEISPHVLTANTIOETAVRAG-----REV 816

QY 301 IMGTLLRRGQG 310
 Db 817 ALGTLRRGEG 826

RESULT 11

US-08-804-227C-4
 ; Sequence 4, Application US/08804227C
 ; Patent No. 5876991
 ; GENERAL INFORMATION:
 ; APPLICANT: Dehoff, Bradley S.
 ; APPLICANT: Kustoss, Stuart A.
 ; APPLICANT: Rostock, Paul R., Jr.
 ; APPLICANT: Sutton, Kimberly L.
 ; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: THOMAS G. PLANT 1501
 ; STREET: LILLY CORPORATE CENTER
 ; CITY: INDIANAPOLIS
 ; STATE: IN
 ; COUNTRY: USA
 ; ZIP: 46285
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: ASCII(DOS) Text only
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/804,227C
 ; FILING DATE: February 21, 1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Plant, Thomas G.
 ; REGISTRATION NUMBER: 35,784
 ; REFERENCE/DOCKET NUMBER: X-8231
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 317-276-2459
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3729 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-804-227C-4

Query Match 47.1%; Score 762; DB 2; Length 3729;
 Best Local Similarity 50.9%; Pred. No. 7.8e-66;
 Matches 161; Conservative 46; Mismatches 99; Indels 10; Gaps 4;

QY 1 VFVPPGSGQWIGMGRQLLETSEVFAAASMDCADALAPHLDMSLDVLRNAAGAHLDHD 60
 Db 2096 VFVPPGSGQWIGMGRQLLETSEVFAAASMDCADALAPHLDMSLDVLRNAAGAHLDHD 2155
 QY 61 DVVQPTLFAMWISLAALMRSGVPCAVLGHSLGEIAAAHVSGLSLADARVVTLSQA 120
 Db 2156 DVVQPALFALWISLAELMRSGVPRVAVVGHSGEIAAACVAGALSVDAARVAVVSRLL 2215
 QY 121 QTTLAGTALVSATPDELPRIPAPWTEENPARLAAVAVNGPRSTVSGAREAVADIVA 180
 Db 2216 QTTLAGTALVSATPDELPRIPAPWTEENPARLAAVAVNGPRSTVSGAREAVADIVA 2271
 QY 181 DLTAQVTRTMIIPDVPAHSPMLVAIERVVSGLPIPTPRSRIPFHSVVGRLDTRRL 239
 Db 2272 DLTAQVTRTMIIPDVPAHSPMLVAIERVVSGLPIPTPRSRIPFHSVVGRLDTRRL 2331
 QY 240 DAAYWRMSSTVRFEPARALLLOOGPKTFVMSPHVLTWGLQELAPDLGDTTGTADTV 299
 Db 2332 DAAYWRMSSTVRFEPARALLLOOGPKTFVMSPHVLTWGLQELAPDLGDTTGTADTV 2387
 QY 300 IMGTLLRRGQG 315
 Db 2388 IMGTLLRRGQG 315

STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII (DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1580 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-227C-11

Query Match 44.9%; Score 726; DB 2; Length 1580;
Best Local Similarity 48.7%; Pred. No. 7.9e-63;
Matches 154; Conservative 37; Mismatches 115; Indels 10; Gaps 3;

QY 1 VFVFPGGQPPWPGNGRELLDASDVFRBSVRACEAFAPAYVMSVEQVLRDSPAPGLDRV 60
DB 565 VFVFPGGQPPWPGNGRELLDASDVFRBSVRACEAFAPAYVMSVEQVLRDSPAPGLDRV 624
QY 61 DVVQPTLFPAVMSIALAAMSQGVPCAVLGHSLGEIAAAHVSGSLSDAARVYVTLMSQA 120
DB 625 DVVQPTLFPAVMSIALAAMSQGVPCAVLGHSLGEIAAAHVSGSLSDAARVYVTLMSQA 684
QY 121 QTTLAGTALVSVAATPDELPRIPAEEDNPARLAANAANGPSTVSGAREVADLVA 180
DB 685 WKRLAGNGMLSVAPADRVRELMPEWAE---RMSVAANGPASVTVAGDARALEEFGG 740
QY 181 DLTAAGY-RTRMIPVDVPAHSPLMYAIEERVVSGLLPTPRSPRIPEHSSVTGRLDTRE 239
DB 741 RLBAAGVLRWPLAGVDPAHSPQVEQFRALDLTLGTVRPAARLPFFSVTAAAEPEG 800
QY 240 LDAAYWRNNSSTVRFEPPARLLLOQGPKEFVMSPHVLTMLGLOELAPDLGDTTGTADT 299
DB 801 LDAAYWRNNSSTVRFEPPARLLLOQGPKEFVMSPHVLTMLGLOELAPDLGDTTGTADT 855
QY 300 VINGTLRRGGCTLDHF 315
DB 856 TALATLHRGSGGLDRF 871

RESULT 15
US-08-804-198-5
Sequence 5, Application US/08804198
Patent No. 5945320
GENERAL INFORMATION:
APPLICANT: Burgett, Stanley G.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rao, Nagataja R.
APPLICANT: Richardson, Mark A.
APPLICANT: Roeseck, Paul R., Jr.
TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS

STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1580 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-5

Query Match 44.9%; Score 726; DB 2; Length 1580;
Best Local Similarity 48.7%; Pred. No. 7.9e-63;
Matches 154; Conservative 37; Mismatches 115; Indels 10; Gaps 3;

QY 1 VFVFPGGQPPWPGNGRELLDASDVFRBSVRACEAFAPAYVMSVEQVLRDSPAPGLDRV 60
DB 565 VFVFPGGQPPWPGNGRELLDASDVFRBSVRACEAFAPAYVMSVEQVLRDSPAPGLDRV 624
QY 61 DVVQPTLFPAVMSIALAAMSQGVPCAVLGHSLGEIAAAHVSGSLSDAARVYVTLMSQA 120
DB 625 DVVQPTLFPAVMSIALAAMSQGVPCAVLGHSLGEIAAAHVSGSLSDAARVYVTLMSQA 684
QY 121 QTTLAGTALVSVAATPDELPRIPAEEDNPARLAANAANGPSTVSGAREVADLVA 180
DB 685 WKRLAGNGMLSVAPADRVRELMPEWAE---RMSVAANGPASVTVAGDARALEEFGG 740
QY 181 DLTAAGY-RTRMIPVDVPAHSPLMYAIEERVVSGLLPTPRSPRIPEHSSVTGRLDTRE 239
DB 741 RLBAAGVLRWPLAGVDPAHSPQVEQFRALDLTLGTVRPAARLPFFSVTAAAEPEG 800
QY 240 LDAAYWRNNSSTVRFEPPARLLLOQGPKEFVMSPHVLTMLGLOELAPDLGDTTGTADT 299
DB 801 LDAAYWRNNSSTVRFEPPARLLLOQGPKEFVMSPHVLTMLGLOELAPDLGDTTGTADT 855
QY 300 VINGTLRRGGCTLDHF 315
DB 856 TALATLHRGSGGLDRF 871

Search completed: June 17, 2003, 13:09:24
Job time : 6.74347 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:50:57 ; Search time 15.48 Seconds
(without alignments)
2720.102 million cell updates/sec

Title: US-09-914-286-3_COPY_29_344

Perfect score: 1618
Sequence: 1 VFVPCGPGQPMGRELID.....ADTVIMGLTRGQGLDHLFL 316

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: A_Geneseq_101002.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1618	100.0	3972	21	AA823749
2	1618	100.0	3972	22	AA65264
3	1618	100.0	3972	22	AA65268
4	819	50.6	344	20	AA87714
5	819	50.6	344	21	AA812926
6	809.5	50.0	3724	18	AAW23718
7	809.5	50.0	3724	18	AAW2608
8	791	48.9	4881	21	AA823751
9	791	48.9	4881	22	AA65266
10	776.5	48.0	3567	14	AA84431

11	765.5	47.3	3192	22	AA810128	Streptomyces nours
12	765	47.3	2595	20	AA539297	Spn a polyketide
13	765	47.3	2595	22	AA570965	S. spinosa protein
14	762	47.1	3729	18	AAW2603	Tylosone synthase
15	753	46.5	7068	22	AA810142	Streptomyces nours
16	753	46.5	9477	22	AA810144	Streptomyces nours
17	750	46.4	5588	20	AA539301	Spn a polyketide
18	750	46.4	5588	22	AA570969	S. spinosa protein
19	746	46.1	3562	22	AA82213	Polyketide synthase
20	734	45.4	5532	21	AA823752	S. avermitilis ave
21	734	45.4	5532	22	AA65267	Streptomyces averm
22	726	44.9	1580	18	AAW23719	Platenolide synth
23	726	44.9	1580	18	AAW2609	Platenolide synth
24	718.5	44.4	3201	22	AA882214	Polyketide synthas
25	714.5	44.2	3816	21	AA592708	S. antibioticus 8,
26	713	44.1	6239	21	AA823750	S. avermitilis ave
27	713	44.1	6239	22	AA65265	Streptomyces averm
28	712	44.0	4150	21	AA592707	S. antibioticus 8,
29	707	43.7	3519	21	AA592709	S. antibioticus 8,
30	704	43.5	3546	22	AA882212	Polyketide synthas
31	702.5	43.4	2986	14	AA844432	eryA region polype
32	698	43.1	3398	14	AA844430	eryA region polype
33	697.5	43.1	5069	19	AAW52846	A. mediterranei tl
34	696.5	43.0	4551	21	AA818637	Amino acid sequenc
35	696.5	43.0	4551	21	AA67201	Narboxolide synth
36	696.5	43.0	4613	21	AA577192	S. venezuelae macr
37	696.5	43.0	4613	21	AA577200	S. venezuelae macr
38	696.5	43.0	12189	21	AA577180	S. venezuelae pik
39	682	42.2	383	23	AA577689	Actinomycete monen
40	681.5	42.1	3739	21	AA818638	Amino acid sequenc
41	681.5	42.1	3739	21	AA577193	S. venezuelae macr
42	681.5	42.1	3739	21	AA577201	S. venezuelae pik
43	681.5	42.1	3739	21	AA67202	Narboxolide synth
44	674.5	41.7	3170	20	AA539299	Spn a polyketide
45	674.5	41.7	3170	22	AA570967	S. spinosa protein

ALIGNMENTS

RESULT 1
AAB23749
AAB23749 standard; Protein; 3972 AA.
XX
AC AAB23749;
XX
DT 10-JAN-2001 (first entry)
XX
DE S. avermitilis avermectin aglycon synthase protein SEQ ID NO:3.
XX
KW Streptomyces avermitilis; avermectin aglycon synthase; biosynthesis;
KW multifunctional enzyme; polyketide; avermectin; veterinary drug;
KW agrochemical.
XX
OS Streptomyces avermitilis.
XX
PN MO200050605-A1.
XX
PD 31-AUG-2000.
XX
PF 23-FEB-2000; 2000MO-JP01041.
XX
PR 24-FEB-1999; 99JP-0046961.
XX
FA (KITA) KITASATO INST.
XX
PI Omura S, Ikeda H;
XX
DR WPI: 2000-565458/52.
XX
DR N-PSDB; AAA82301.
XX
PT Avermectin aglycone synthase DNA and proteins encoded by all or part of
it for the production of avermectin and its derivatives for drug and

PT agrochemical use -
XX
XX Claim 32; Page 203-223; 314pp; Japanese.
XX
CC The present sequence represents an avermectin aglycon synthase protein.
CC Also described are: (1) polypeptides encoded by all or part of the DNA;
CC (2) expression vectors containing the DNA; (3) host cells transformed by
CC the vectors; (4) preparation of the polypeptides by culture of the
CC transformants; (5) preparation of avermectin aglycon or its derivatives
CC by culture of transformed avermectin-producing microorganisms; and (6)
CC oligonucleotides of 5-60 bases in length containing sense or antisense
CC sequences from the avermectin aglycon synthase DNA. The enzymes are
CC useful for the production of modified forms of avermectin and of the
CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
CC and agrochemicals.
XX
SQ Sequence 3972 AA;
Query Match 100.0%; Score 1618; DB 21; Length 3972;
Best Local Similarity 100.0%; Pred. No. 3.2e-138;
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VFVPPGGQPPQWPGMREILDASDVFRBSVACBAAPAPYVDWVVEQVLRDSPAPGLDRV 60
DB 29 VFVPPGGQPPQWPGMREILDASDVFRBSVACBAAPAPYVDWVVEQVLRDSPAPGLDRV 88
QY 61 DVVQPTLFAMWISIALALMRSQGVPCAVLGHSLGEIAAAHVSGSLSLADAAVVTLMSQA 120
DB 89 DVVQPTLFAMWISIALALMRSQGVPCAVLGHSLGEIAAAHVSGSLSLADAAVVTLMSQA 148
QY 121 QTTLAGTALVSAATPDELLPRIAPWTEDNPARLAVAANGRSTVSGAREAVADLVA 180
DB 149 QTTLAGTALVSAATPDELLPRIAPWTEDNPARLAVAANGRSTVSGAREAVADLVA 208
QY 181 DLTAAQVTRTMIIPVDPAHSPLMTVAIEERVVSGLLPTTPRPSRIPFHSSVTGGRLDTREL 240
DB 209 DLTAAQVTRTMIIPVDPAHSPLMTVAIEERVVSGLLPTTPRPSRIPFHSSVTGGRLDTREL 268
QY 241 DAAVWYRNMSSTVRFEPARALLQGGPKTFVEMSPHVLTMGLQELAPDLGDTTGTADTV 300
DB 269 DAAVWYRNMSSTVRFEPARALLQGGPKTFVEMSPHVLTMGLQELAPDLGDTTGTADTV 328
QY 301 IMGTLRGGQGLDHFLL 316
DB 329 IMGTLRGGQGLDHFLL 344
RESULT 2
AAG65264
ID AAG65264 standard; Protein; 3972 AA.
XX
AC AAG65264;
XX
DT 04-DEC-2001 (first entry)
XX
DE Streptomycetes avermilitilis protein SEQ ID NO: 4.
XX
KM Avermectin aglycone synthase; AAS; avermectin derivative;
KM drug production; veterinary drug; pesticide.
XX
OS Streptomycetes avermilitilis.
XX
OS WO200162939-A1.
XX
PN 30-AUG-2001.
XX
PD 23-FEB-2001; 2001WO-JP01381.
XX
PF 24-FEB-2000; 2000JP-0047405.
XX
PR (KYOW) KYOWA HAKKO KOGYO KK.
PA (KITA) KITASATO INST.
XX

PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
XX
XX WPI; 2001-582053/65.
DR N-PSDB; AAH79277.
XX
XX New modified avermectin aglycone synthase derived from Streptomycetes
PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
PT drugs and pesticides -
XX
XX Claim 4; Page 167-180; 257pp; Japanese.
XX
XX The present invention relates to the production of modified derivatives
CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
CC avermilitilis. The activity of an acyl carrier protein (ACP),
CC beta-ketoadyl carrier protein synthase (KS), acyltransferase (AT),
CC beta-ketoadyl carrier protein reductase (KR), dehydratase (DH), enoyl
CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC suppressed. The process can be used in the production of drugs, veterinary
CC drugs and pesticides. The present sequence is an S. avermilitilis protein.
XX
SQ Sequence 3972 AA;
Query Match 100.0%; Score 1618; DB 22; Length 3972;
Best Local Similarity 100.0%; Pred. No. 3.2e-138;
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VFVPPGGQPPQWPGMREILDASDVFRBSVACBAAPAPYVDWVVEQVLRDSPAPGLDRV 60
DB 29 VFVPPGGQPPQWPGMREILDASDVFRBSVACBAAPAPYVDWVVEQVLRDSPAPGLDRV 88
QY 61 DVVQPTLFAMWISIALALMRSQGVPCAVLGHSLGEIAAAHVSGSLSLADAAVVTLMSQA 120
DB 89 DVVQPTLFAMWISIALALMRSQGVPCAVLGHSLGEIAAAHVSGSLSLADAAVVTLMSQA 148
QY 121 QTTLAGTALVSAATPDELLPRIAPWTEDNPARLAVAANGRSTVSGAREAVADLVA 180
DB 149 QTTLAGTALVSAATPDELLPRIAPWTEDNPARLAVAANGRSTVSGAREAVADLVA 208
QY 181 DLTAAQVTRTMIIPVDPAHSPLMTVAIEERVVSGLLPTTPRPSRIPFHSSVTGGRLDTREL 240
DB 209 DLTAAQVTRTMIIPVDPAHSPLMTVAIEERVVSGLLPTTPRPSRIPFHSSVTGGRLDTREL 268
QY 241 DAAVWYRNMSSTVRFEPARALLQGGPKTFVEMSPHVLTMGLQELAPDLGDTTGTADTV 300
DB 269 DAAVWYRNMSSTVRFEPARALLQGGPKTFVEMSPHVLTMGLQELAPDLGDTTGTADTV 328
QY 301 IMGTLRGGQGLDHFLL 316
DB 329 IMGTLRGGQGLDHFLL 344
RESULT 3
AAG65268
ID AAG65268 standard; Protein; 3972 AA.
XX
AC AAG65268;
XX
DT 04-DEC-2001 (first entry)
XX
DE Streptomycetes avermilitilis protein derivative SEQ ID NO: 8.
XX
KM Avermectin aglycone synthase; AAS; avermectin derivative;
KM drug production; veterinary drug; pesticide.
XX
OS Synthetic.
XX
OS WO200162939-A1.
XX
PN 30-AUG-2001.
XX
PD 23-FEB-2001; 2001WO-JP01381.
XX
PF 24-FEB-2000; 2000JP-0047405.
XX

XX (RYOW) KYOMA HAKKO KOGYO KK.
 PA (KITA) KITASATO INST.
 XX
 XX
 PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
 DR WPI, 2001-562053/65.
 DR N-PSDB; AAB79279.
 XX
 XX New modified avermectin aglycone synthase derived from Streptomyces
 PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
 PT drugs and pesticides
 PS Claim 5; Page 235-248; 257pp; Japanese.
 XX
 XX The present invention relates to the production of modified derivatives
 CC of avermectin aglycone synthase (AAS) derived from Streptomyces
 CC avermectilis. The activity of an acyl carrier protein (ACP),
 CC beta-ketoadacyl carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketoadacyl carrier protein reductase (KR), dehydratase (DH), enoyl
 CC reductase (ER) and/or cholesterylase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is derived from an S.
 CC avermectilis protein.
 CC
 SQ Sequence 3972 AA;
 XX
 XX Query Match 100.0%; Score 1618; DB 22; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 3.2e-138;
 Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VFVFPQGGPQWPGMGRELLDASDVFRSVACAEAFAPYVDWSVEQVLRDSPAPGLDRV 60
 DB 29 VFVFPQGGPQWPGMGRELLDASDVFRSVACAEAFAPYVDWSVEQVLRDSPAPGLDRV 88
 QY 61 DVVQPTLFAMVMSLAALMRSQGVPCAVLGHSLGEIAAAVSGGLSLADAAVYVTTMSQA 120
 DB 89 DVVQPTLFAMVMSLAALMRSQGVPCAVLGHSLGEIAAAVSGGLSLADAAVYVTTMSQA 148
 QY 121 QTTLAGTALVSAATPDELPRIPAPWTEEDNPRLAAVAVNGRSTVVGAREAVADVLA 180
 DB 149 QTTLAGTALVSAATPDELPRIPAPWTEEDNPRLAAVAVNGRSTVVGAREAVADVLA 208
 QY 181 DLTAQVTRMIPVDVPAHSPLMYAIEERVVSGLLPTTPRSRIIPFHSVYTGRLDTR 240
 DB 209 DLTAQVTRMIPVDVPAHSPLMYAIEERVVSGLLPTTPRSRIIPFHSVYTGRLDTR 268
 QY 241 DAAVYTRNMSSTVFEPAARLLIQGPXTVEMSPHVLTMGLQELAPDLGDTTGTADTV 300
 DB 269 DAAVYTRNMSSTVFEPAARLLIQGPXTVEMSPHVLTMGLQELAPDLGDTTGTADTV 328
 QY 301 IMGTLRGOGTLDHFL 316
 DB 329 IMGTLRGOGTLDHFL 344
 XX
 XX RESULT 4
 AAB87714
 ID AAB87714 standard; Protein; 344 AA.
 XX
 AC AAB87714;
 XX
 DT 09-MAR-1999 (first entry)
 XX
 DE An active acyltransferase domain (NidA75) from a polyketide synthase.
 XX
 XX Enzymatically active acyltransferase domain; polyketide synthase; PKS;
 KW polyketide synthase; antimicrobial; anticancer; antifungal;
 KM immunosuppressant; antihelminthic; NidA75; ethyl acyltransferase domain;
 KM module 5; PKS cluster.
 XX
 OS Streptomyces caelestis.
 XX

PN WO9851695-A2.
 XX
 XX 19-NOV-1998.
 PD
 XX
 XX 13-MAY-1998; 98WO-US09518.
 PF
 XX 16-MAY-1997; 97US-0858003.
 PR
 XX
 XX (ABBO) ABBOTT LAB.
 PA
 PI Kakaya SJ, Katz L, Pereda-Lopez A, Ruan X, Staasi DL;
 PI Summers RG;
 PI
 DR WPI; 1999-070114/06.
 DR N-PSDB; AAV9255.
 XX
 XX
 XX New polyketides produced by microorganism having altered polyketide
 PT synthase genes - especially chimeras containing a heterologous
 PT acyltransferase domain, useful as e.g. antimicrobials, anticancer
 PT agents
 PS Claim 25; Fig 21; 137pp; English.
 XX
 XX The present sequence represents an enzymatically active acyltransferase
 CC domain, NidA75, from a polyketide synthase (PKS). The NidA75 protein
 CC is derived from the ethyl acyltransferase domain from module 5 of the
 CC PKS cluster of Streptomyces caelestis NRRL-2821. The sequence
 CC can be used to replace acyltransferase domains of other PKS enzymes,
 CC which are then used to generate novel polyketides by de novo biosynthesis
 CC rather than by chemical modification. Polyketides of the invention have
 CC the methyl groups on the macrocyclic ring substituted with -H, -Et
 CC and/or -OH, and the ethyl side chain substituted with hydroxymethyl or
 CC dihydroxycyclohexylmethyl side chains. Polyketides of the invention are
 CC antimicrobial, anticancer, antifungal, immunosuppressant and/or
 CC antihelminthic agents, or their intermediates.
 CC
 SQ Sequence 344 AA;
 XX
 XX Query Match 50.6%; Score 819; DB 20; Length 344;
 Best Local Similarity 54.0%; Pred. No. 1.5e-66;
 Matches 168; Conservative 38; Mismatches 95; Indels 10; Gaps 3;
 QY 1 VFVFPQGGPQWPGMGRELLDASDVFRSVACAEAFAPYVDWSVEQVLRDSPAPGLDRV 60
 DB 5 VFVFPQGGPQWPGMGRELLDASDVFRSVACAEAFAPYVDWSVEQVLRDSPAPGLDRV 64
 QY 61 DVVQPTLFAMVMSLAALMRSQGVPCAVLGHSLGEIAAAVSGGLSLADAAVYVTTMSQA 120
 DB 65 DVVQPTLFAMVMSLAALMRSQGVPCAVLGHSLGEIAAAVSGGLSLADAAVYVTTMSQA 124
 QY 121 QTTLAGTALVSAATPDELPRIPAPWTEEDNPRLAAVAVNGRSTVVGAREAVADVLA 180
 DB 125 WTLGAGKGMVAAPPAEELRPLVTWGD---RLAAVAVNSGSCAVAGDPALAEVLA 180
 QY 181 DLTAQVTRMIPVDVPAHSPLMYAIEERVVSGLLPTTPRSRIIPFHSVYTGRLDTR 239
 DB 181 LITGEGVHARPIPGVDTAGHSQVDLRAHLEVLAVAPRAPRADIIFYETVYTGRLDTR 240
 QY 240 LDAVYTRNMSSTVFEPAARLLIQGPXTVEMSPHVLTMGLQELAPDLGDTTGTADTV 299
 DB 241 LDAVYTRNMSSTVFEPAARLLIQGPXTVEMSPHVLTMGLQELAPDLGDTTGTADTV 295
 QY 300 VIMGTLRGOG 310
 DB 296 AVLGTLRRHG 306
 XX
 XX RESULT 5
 AAB12926
 ID AAB12926 standard; Protein; 344 AA.
 XX
 AC AAB12926;
 XX

DT	20-NOV-2000	(first entry)
XX		
DE	Protein sequence of the ethyl AT module of the niddamycin PKS cluster.	
XX		
KW	Erythromycin derivative; polyketide synthase; PKS; acetyltransferase;	
KW	antibiotic; antifungal; anticancer; immunosuppressant; antihelminthic;	
XX	ethyl AT; niddamycin.	
XX		
OS	Streptomyces caelestis.	
XX		
PN	US6063561-A.	
XX		
PD	16-MAY-2000.	
XX		
PF	13-MAY-1998; 98US-0078166.	
XX		
PR	17-JAN-1991; 91US-0642734.	
PR	16-MAY-1997; 97US-0858003.	
XX		
PA	(ABBO) ABBOTT LAB.	
XX		
P1	Katz U, Stasael DL, Ruan X, Pereda-Lopez A, Kakavas SU, Summers RG;	
DR	WPI; 2000-386760/33.	
DR	N-PSDB; AAA72521.	
XX		
FT	New erythromycin derivatives produced by microorganisms with altered	
PT	polyketide synthase genes -	
XX		
PS	Example 20; Fig 21; 94pp; English.	
CC	This invention relates to novel erythromycin derivatives, in which the	
CC	methyl groups on the macro lactone ring have been substituted with H, Et	
CC	and/or OH and the ethyl side chain has been substituted with	
CC	hydroxymethyl or dihydroxycyclohexylmethyl side chains. Included in the	
CC	invention are vectors containing polynucleotide sequences which encode an	
CC	enzymatically active acetyltransferase domain from Streptomyces. Also	
CC	included are methods for altering the substrate specificity of a	
CC	polyketide synthase in a microorganism using the vectors of the	
CC	invention. Polyketides are are a large class of natural products	
CC	including antibiotics, antifungal, anticancer and immunosuppressant	
CC	compounds. Polyketide synthase (PKS) is a complex of enzyme activities	
CC	responsible for the biosynthesis of polyketides. The erythromycin	
CC	derivatives exhibit antibiotic, antifungal, anticancer, antihelminthic	
CC	and immunosuppressant activity.	
CC	The present sequence represents the amino acid sequence of the ethyl	
CC	acetyltransferase (AT) module of the niddamycin PKS cluster isolated from	
CC	Streptomyces caelestis. The sequence is used in the production of a	
CC	vector of the invention.	
XX		
SO	Sequence 344 AA;	
	Query Match 50.6%; Score 819; DB 21; Length 344;	
	Best Local Similarity 54.0%; Pred. No. 1.5e-66;	
	Matches 168; Conservativity 38; Mismatches 95; Indels 10; Gaps 3	
QY	1 VFVPPGGQPPWPGMGRELLIDASDVRESVRACEAFAPYVDMVSVQVLRSDPADGLDRV 60	
DB	5 VFVPPGGQSQWAGNAAGELERSGARFSAADSCDALRPLGMSVLSVLRGEPDASLDRV 64	
QY	61 DVVQPTLFAVMTSLAALWRSQGVCECAVLGSHSGEIAAAHVSGGSLADAARVTLMSQA 120	
DB	65 DVVQGPVLTETMTVMVSLAAVWRALGVESPAVAVGSHSGEIAAAHVAAGALSILDSARIVALRSRA 124	
QY	121 QTTLAGTCALVSVAATPDELLPRIAPWTFEDNPARIAAVAAYVGPSTVSGAREAVADIYA 180	
DB	125 WLGLGKGGKGMVAVPMFABELRDLRTWGD---RLAAVAAVNSPGSCAAVAGDEALAEIYA 180	
QY	181 DLTAAGVATRMIP-VDVPAHSEPLMTVAIEERVVSGLPTTPRPSRIPFHSSVTGRLDRE 239	
DB	181 LLTGSGVHARPIPGVDTGASHSQVDALRAHLLEVLVAFAVPRADIPFYSTVTGGLDSTE 240	
QY	240 LDAAVYRNMSSTVFEEPARILLLOQGPRTVEMSPPHVLTMGLQELALDGLDGTGTADT 299	

Db	241	LDATYRNNRREVEERERATRLADGHDVLEISPHMLVALEQVTDAG-----TDA	295
Oy	300	VINGTLRRGG 310	
Db	296	AVIGTLRRRG 306	
RESULT 6			
ID	AAW23718	standard; Protein; 3724 AA.	
AC	AAW23718;		
XX	27-FEB-1998	(first entry)	
XX		Platenolide synthase ORF3 protein.	
DE		Platenolide synthase gene cluster; platenolide production; ermG gene;	
XX		multi-functional protein; macrolide antibiotic; spiramycin.	
KW		Streptomyces ambofaciens.	
XX			
OS			
XX			
XX	Key	Location/Qualifiers	
FT	Modified-site	1	
FT		/note= "encoded by TGT"	
FT	Domain	36..459	
FT		/note= "ketosynthase domain, KS4"	
FT	Domain	594..921	
FT		/note= "acyltransferase domain, Atr4"	
FT	Domain	1177..1350	
FT		/note= "ketosynthase domain, KS04"	
FT	Domain	1459..1523	
FT		/note= "acyl carrier protein domain, ACP4"	
FT	Domain	1545..1969	
FT		/note= "ketosynthase domain, KS5"	
FT	Domain	2088..2424	
FT		/note= "acyltransferase domain, ATr5"	
FT	Domain	2445..2619	
FT		/note= "dehydratase domain, DHS"	
FT	Domain	2912..3261	
FT		/note= "enoylreductase domain, ER5"	
FT	Domain	3266..3451	
FT		/note= "ketoreductase domain, KR5"	
FT	Domain	3551..3631	
FT		/note= "acyl carrier protein domain, ACP5"	
XX			
PV	EP791656-A2.		
XX			
XX	27-AUG-1997.		
PD			
XX	19-FEB-1997;	97EP-0301066.	
PF			
XX	22-FEB-1996;	96US-0012050.	
PR			
XX			
PA	(EHL) LILLY & CO ELI.		
PI	Burgelt SG, Kuhstose SA, Rao RN, Richardson MA;		
PI	Rosteck PR;		
XX			
XX	WPI; 1997-418047/39.		
DR	N-PSDB; AAT78508.		
XX			
XX	DNA encoding Streptomyces ambofaciens platenolide synthase domain -		
PT	for production of spiramycin-related polyketide antibiotics		
XX			
XX	Claim 8, Pages 54-66; 81pp; English.		
XX			
CC	AAW23716-W23720	represent proteins encoded by the platenolide synthase	
CC		gene cluster of the invention. The gene cluster is also referred to as	
CC		the ermG gene, and was isolated from Streptomyces ambofaciens. These	
CC		sequences are multi-functional proteins which direct the synthesis of	
CC		the polyketide platenolide. Platenolide is the basic building block of	

RESULT 8

ID AAB23751 standard; Protein; 4881 AA.

XX AAB23751;

XX 10-JAN-2001 (first entry)

XX S. avermiltis avermectin aglycon synthase protein SEQ ID NO:5.

XX Streptomyces avermiltis; avermectin aglycon synthase; biosynthesis;
XX multifunctional enzyme; polyketide; avermectin; veterinary drug;
XX agrochemical.

XX Streptomyces avermiltis.

XX W0200050605-A1.

XX 31-AUG-2000.

XX 23-FEB-2000; 2000WO-JP01041.

XX 24-FEB-1999; 99JP-0046961.

XX (KITA) KITASATO INST.

XX Omura S, Ikeda H;

XX WPI; 2000-565458/52.

XX N-PSDB; AAA92302.

XX Avermectin aglycone synthase DNA and proteins encoded by all or part of
PT it for the production of avermectin and its derivatives for drug and
PT agrochemical use

XX Claim 32; Page 254-278; 314pp; Japanese.

XX The present sequence represents an avermectin aglycon synthase protein.
 CC Also described are: (1) polypeptides encoded by all or part of the DNA;
 CC (2) expression vectors containing the DNA; (3) host cells transformed by
 CC the vectors; (4) preparation of the polypeptides by culture of the
 CC transformants; (5) preparation of avermectin aglycon or its derivatives
 CC by culture of transformed avermectin-producing microorganisms; and (6)
 CC oligonucleotides of 5-60 bases in length containing sense or antisense
 CC sequences from the avermectin aglycon synthase DNA. The enzymes are
 CC useful for the production of modified forms of avermectin and of the
 CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
 CC and agrochemicals.

XX Sequence 4881 AA;

XX Query Match 48.9%; Score 791; DB 21; Length 4881;

XX Best Local Similarity 52.2%; Pred. No. 1.7e-62;

XX Matches 164; Conservative 45; Mismatches 97; Indels 8; Gaps 2;

XX 3 VFPQGGQWMPGMRRELLDASVFRSEVRACAAFAFPYVDSVEQVLRDSPAGLDRVDV 62

XX 2139 LFGQGGTHWGMQDELLGSSPVFAQMSDCQALBEPYLDMLDLVLRGADAPPLQGRVDV 2198

XX 63 VQPTLFAVMISLAIAWRSQGVPCAVLGHSIGETIAAAHVSGSLADAAARVTLMSQAOT 122

XX 2199 VQPTLFAVMISLAIAWRSQGVPCAVLGHSIGETIAAAHVSGSLADAAARVTLMSQAOT 2258

XX 123 TLAAGTALVSVAAATPDELPRIPAWTEBDNPARLAVAANGPRSTVSGAREAVADLVADL 182

XX 2259 ALAAGGAMASVGLPEVKLEPRLATWGD---RLVIAAVNGARSAVSGEPEAVDALVEEL 2314

XX 183 TAAQVRIRMTFVDVPAISPLMTYALBEERVVSGLLPITRPRSPITPHSSVTGRLDTREIDA 242

XX 2315 SHEDVPAARLWVAVASHSPQVEALQGRILLELPIARATGDPVPSVTGTGERIDGTBLDA 2374

XX 243 AYWRNMSSTVREPPARLLLOOGPKTFVEMSPHVLTMGLQELAPDLGDTGTADTVIM 302

XX 2375 DMYRNRNRQVRRFADATQALVRAGHTVFTEACHPAIVAGVGTLDLMDGD----LDSLVV 2430

XX 303 GTLRGGQGTLDHFL 316

XX 2431 GSLRRGEGGLRFL 2444

RESULT 9

ID AAG65266 standard; Protein; 4881 AA.

XX AAG65266;

XX 04-DEC-2001 (first entry)

XX Streptomyces avermiltis protein SEQ ID NO: 6.

XX Avermectin aglycone synthase; AAS; avermectin derivative;

XX drug production; veterinary drug; pesticide.

XX Streptomyces avermiltis.

XX W0200162939-A1.

XX 30-AUG-2001.

XX 23-FEB-2001; 2001WO-JP01381.

XX 24-FEB-2000; 2000JP-0047405.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX (KITA) KITASATO INST.

XX Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;

XX WPI; 2001-582053/65.

XX N-PSDB; AAH79278.

XX New modified avermectin aglycone synthase derived from Streptomyces
PT avermectin used in production of 22,23-dihydroavermectin B1a used in
PT drugs and pesticides

XX Claim 4; Page 201-217; 257pp; Japanese.

XX The present invention relates to the production of modified derivatives
 CC of avermectin aglycone synthase (AAS) derived from Streptomyces
 CC avermiltis. The activity of an acyl carrier protein (ACP),
 CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
 CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is an S. avermiltis protein.

XX Sequence 4881 AA;

XX Query Match 48.9%; Score 791; DB 22; Length 4881;

XX Best Local Similarity 52.2%; Pred. No. 1.7e-62;

XX Matches 164; Conservative 45; Mismatches 97; Indels 8; Gaps 2;

XX 3 VFPQGGQWMPGMRRELLDASVFRSEVRACAAFAFPYVDSVEQVLRDSPAGLDRVDV 62

XX 2139 LFGQGGTHWGMQDELLGSSPVFAQMSDCQALBEPYLDMLDLVLRGADAPPLQGRVDV 2198

XX 63 VQPTLFAVMISLAIAWRSQGVPCAVLGHSIGETIAAAHVSGSLADAAARVTLMSQAOT 122

XX 2199 VQPTLFAVMISLAIAWRSQGVPCAVLGHSIGETIAAAHVSGSLADAAARVTLMSQAOT 2258

XX 123 TLAAGTALVSVAAATPDELPRIPAWTEBDNPARLAVAANGPRSTVSGAREAVADLVADL 182

XX 2259 ALAAGGAMASVGLPEVKLEPRLATWGD---RLVIAAVNGARSAVSGEPEAVDALVEEL 2314

```

Oy      183   AAAGVRTMIPDVDAHSPWVAIEBRYVSGLLPTTPRPSRIPEFSSVTGGRLDREIDA 242
Db      2315  SHEVDPAARLAWMDASHSBOYEAIGCRLLIELAPARTGDPVPFSITVGEEIDGTLEDA 2376
Oy      243   AYWVENMSYTRFEEDPARULLQOGKPTFVENSPPHFVLTMGLQELAPDUGDTTGTADTVIM 302
Db      2375  DYWRNINLVQRFRATQALVRAHTVFIEACPFAVAVAVGVETLDEMGD-----LDSLVV 2433
Oy      303   GTLRRGQGTLDHFL 316
Db      2431  GSLRREGGLRRFL 2444

RESULT 10
AAR44431
ID      AAR44431 standard; Protein; 3567 AA.
AC      AAR44431/
XX      22-DEC-1993 (first entry)
Dt      eryA region polypeptide module #2.
De      Saccaropolyspora erythraea; eryA; biosynthesis; polyketide; module;
Xx      erychromycin; condensation; elongation; acyl chain growth;
Km      gene replacement.
Xk      Saccaropolyspora erythraea.
Os      MO9J13663-A.
Pn      22-JUL-1993.
Pd      17-JAN-1992; 92WO-US00427.
Pe      17-JAN-1992; 92MO-US00427.
Pr      17-JAN-1992; 92MO-US00427.
Px      (ABBO ) ABBOTT LAB.
Py      Donadio S, Katz L, McAlpine JB;
Pz      MPI, 1993-242804/30.
Dr      N-P8DB; AAA046806.
Dx      Biosynthesis of specific polyketide analogues esp. erythromycin
Pt      cpds. - by introducing altered biosynthetic gene-cong. DNA into
Pv      microorganisms
Ps      Disclosure; Fig 2, 133pp; English.
Pz      The sequences given in AAR44430-32 are encoded by the eryA fragment of
xx      the Saccarapolyspora erythraea genome.. These polypeptides are
cc      involved in the biosynthesis of the polyketide segment of erythromycin.
cc      eryA is organised in modules and each module takes care of one
cc      condensation step. The precise succession of elongation steps is
cc      dictated by the genetic order of the modules. The DNA encoding
cc      these polypeptides may be specifically altered such that novel
cc      polyketide molecules of desired structure are produced. Three types
cc      of alteration may be produced: those inactivating a single function in
cc      a module which does not arrest acyl chain growth; those inactivating a
cc      single function in a module which does affect chain growth; and those
cc      affecting an entire module. The mutations may be introduced by gene
cc      replacement.
xx      xx
Sd      Sequence 3567 AA;

Query Match 48.0%; Score 776.5; DB 14; Length 3567;
Best Local Similarity 50.0%; Pred. No. 2,3e-61;
Matches 158; Conservative 44; Mismatches 107; Indels 7; Gaps 2;

1 VFVPFGGQPMWGMRRELILDASDVRESVRACEAAFPAPVVMGSVQVLRDSPADARLDRV 60
560 VFLEPGGQSOMYGKAELISSPVAGKRIRACDSMAPQDKVSDVLRQAAGAGLDRV 619

```

Oy		61 VVVOPTLPAWMIISLAALNRSGVECAVLGSHLGETAIAAHVSGGLSLADAAARVVTLMSQ 120
Dd		620 DVQDPVFAPAWMSIALBELMRSYGVPEAAVAVHSGSEIAAHHAGALTLEDNAKLIVGRSRL 679
Oy		121 QTLTAGTALVSVAATPDELLPRILPWTEEDNPARLAVAANVPSTSVSGAREAVADLVA 180
Dd		680 MRSLSGECSMAVALGAEAVERERLPMPD---RLSVAANVPSTSVSGEPCALRAFSE 735
Oy		181 DLTAQVTRIMIPVDVPAHSPLMTVIIEERVVSGLLPITPRPERIPRHSSVTGRLLTREI 240
Dd		736 DCALEGIRVRVIDDVAYASHSPDIKVERRELETTGGDILARPAPRVFHSITVESRSMGTET 795
Oy		241 DAAYWRMSSTVREPAPARLLLOQSPTPVEMSPPHPLYTMLQLSLAPLDGDTTGADTV 300
Dd		796 DARVYRRRLRETVPADAVTRLAESGYDAFIIVESHPPVVQAVERA---VEPADGAEADV 852
Oy		301 IMGTLRNGOFTLIDHFL 316
Dd		853 VVGSLHRDGDLSAFL 868
 RESULT 11 AAE10128 AAE10128 standard; Protein; 3192 AA.		
XX	AC	AAE10128;
XX	XX	
XX	XX	29-NOV-2001 (first entry)
De	St	Streptomyces noursei nystatin gene, NysB.
De	XX	
Kw	Pk	Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster; antifungal; antibiotic; PKS type I.
XX	XX	
OS	Str	Streptomyces noursei.
XX	XX	
Ft	Key	Location/Qualifiers
Ft	Ft	Domain
Ft	Ft	/label= KS1_domain
Ft	Ft	/note= "Ketosynthase (KS) domain"
Ft	Ft	578..897
Ft	Ft	/label= AT1_domain
Ft	Ft	/note= "Acyltransferase (AT) domain"
Ft	Ft	911..1110
Ft	Ft	/label= DH1_domain
Ft	Ft	/note= "Dehydratase (DH) domain"
Ft	Ft	1201..1447
Ft	Ft	/label= KR1_domain
Ft	Ft	/note= "Ketoreductase (KR) domain"
Ft	Ft	1484..1559
Ft	Ft	/label= ACP1_domain
Ft	Ft	/note= "Acyl carrier protein (ACP) domain"
Ft	Ft	1579..2004
Ft	Ft	/label= KS2_domain
Ft	Ft	/note= "Ketosynthase (KS) domain"
Ft	Ft	2117..2439
Ft	Ft	/label= AT2_domain
Ft	Ft	/note= "Acyltransferase (AT) domain"
Ft	Ft	2453..2659
Ft	Ft	/label= DH2_domain
Ft	Ft	/note= "Dehydratase (DH) domain"
Ft	Ft	2749..2996
Ft	Ft	/label= KR2_domain
Ft	Ft	/note= "Ketoreductase (KR) domain"
Ft	Ft	3025..3102
Ft	Ft	/label= ACP2_domain
Ft	Ft	/note= "Acyl carrier protein (ACP) domain"
XX	PN	WO200159126-A2.
XX	PD	16-AUG-2001.
XX	XX	

PF 08-FEB-2001; 2001WO-GB00509.
 XX
 PR 08-FEB-2000; 2000GB-0002840.
 PR 10-APR-2000; 2000GB-0008786.
 PR 14-APR-2000; 2000GB-0009387.
 XX
 PA (UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
 PA (SMTF) SINTEF STIFTELSEN IND TEK FORSK.
 PA (ALPH-) ALPHARMA AS.
 PA (SINV-) SINVENT AS.
 PA (DZIE/) DZIELEMSKA H.
 PA (ZOTC/) ZOTCHEV S B.
 PA (SEKU/) SEKUROVA O N.
 PA (FJAE/) FJAEVVIK E.
 PA (BRAU/) BRAUTASET T.
 PA (STRO/) STROM A R.
 XX
 PI Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR,
 PI Valia S, Ellingsen TE, Sletta H, Gulliksen O;
 XX
 DR WPI: 2001-557614/62.
 DR N-PSDB; AAD17184.
 XX
 PS New nystatin polyketide synthase polynucleotides and polypeptides,
 PT useful as antibiotics and antifungals -
 XX
 PS Claim 15; Page 168-169; 266pp; English.
 CC The present invention relates to the cloning and sequencing of the gene
 CC cluster encoding a modular type I polyketide synthase (PKS) enzyme
 CC involved in the biosynthesis of the macrocyclic antibiotic nystatin.
 CC The nystatin PKS is useful as antifungal antibiotics. The present
 CC sequence is a PKS type I encoding Streptomyces noursei nystatin
 CC gene, NySB.
 XX
 SQ Sequence 3192 AA;
 Query Match 47.3%; Score 765.5; DB 22; Length 3192;
 Best Local Similarity 50.6%; Pred. No. 2e-60;
 Matches 160; Conservative 42; Mismatches 109; Indels 5; Gaps 2;
 QY 1 VFVPPGSGPOMPGNGRELLDASDVFRSVPACFAAPFYDMSVEOVLKSDPAAGLDLV 60
 DB 2114 VFVPPGSGSGMGMGRELATSEVFRATIDCATALAPYDMSVLDVLAGEPDLIERV 2173
 QY 61 DVVQPTLFAYWISLAALMRSGVPCAVLGSLSLEIAAHYSGSLSLADARVYTMQA 120
 DB 2174 DVVQPALFAMWVGLSALMRSHGVPAVAVGSHSGEIAACVAGALSLADARVALRSQA 2233
 QY 121 QTTLAGTALVSVAATPEDELLPRIAPWTEDENPALAAVAVNGPSTVSGAREAVADIVA 180
 DB 2234 LPQLSGSGMWSVSAVERVYALALAPQE---ALSVAAVNGPSSVVGSTDLALDLHT 2289
 QY 181 DLTAQVTRMTIPVDPAHSPMLVAIEERVVSGLLPTTPRESRIPIHSSVTGSLDTREL 240
 DB 2290 ACQEGFRARFVSVDYASHGRHVAEVRDELARVLAVDPRAPEVPFSTVGRDVDAAF 2349
 QY 241 DAAVYVNMSTVFEPAARLLLOQSPKTFVMSPHVLTIMGLOELAPDLGDTTGTADTV 300
 DB 2350 DGAATWYNLRQYVMEBETRLLAAGHRVLEVPHPVLAPIDETGEAFAEATG-GSAV 2408
 QY 301 IMGTLRGQGLDHFLL 316
 DB 2409 VLGSLRRDEGGPRRFL 2424

XX
 DE SpnA a polyketide synthase.
 XX
 KM Spinosyn biosynthetic enzyme; open reading frame; ORF; insecticidal
 KM microtides; arachnid; nematode; insect; polyketide; polyketide synthase;
 KM PKS; extender module; initiator module; acyl transferase domain; AT;
 KM acyl carrier protein; ACP; beta-ketosynthase domain; KS; KR;
 KM dehydratase domain; DH; enoyl reductase domain; ER; beta-ketoreductase;
 KM insecticide.
 XX
 OS Saccharopolyspora spinosa.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1944
 FT /note= "Encoded by CAG"
 FT 6..423
 FT /label= KS1
 FT /note= "beta-ketosynthase domain: part of the
 FT initiator module"
 FT Domain 528..853
 FT /label= AT1
 FT /note= "acyl transferase domain: part of the
 FT initiator module"
 FT Domain 895..977
 FT /label= ACP1
 FT /note= "acyl carrier protein domain: part of the
 FT initiator module"
 FT Domain 998..1413
 FT /label= KS1
 FT /note= "beta-ketosynthase domain: part of extender
 FT module 1"
 FT 1525..1858
 FT /label= AT1
 FT /note= "acyl transferase domain: part of extender
 FT module 1"
 FT Domain 2158..2337
 FT /label= KR1
 FT /note= "beta-ketoreductase domain: part of extender
 FT module 1"
 FT Domain 2432..2513
 FT /label= ACP1
 FT /note= "acyl carrier protein domain: part of extender
 FT module 1"
 XX
 PN WO9946387-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 16-FEB-1999; 99WO-US03212.
 XX
 PR 09-MAR-1998; 98US-0036987.
 XX
 PA (DOWC) DOW AGROSCIENCES LLC.
 PA
 PI Baltz RH, Broughton MC, Crawford KP, Madduri K, Merlo DJ,
 PI Treadway PJ, Turner JR, Waldron C;
 XX
 DR WPI: 1999-551414/46.
 DR N-PSDB; AAZ21501.
 XX
 PT New spinosyn biosynthetic genes from Saccharopolyspora spinosa, useful
 PT for production of insecticidal spinosyn compounds -
 XX
 PS Claim 1; Page 87-96; 190pp; English.
 XX
 CC This is the amino acid sequence of the product of the spnA gene. The
 CC protein is involved in spinosyn biosynthesis. The spnA gene is one of 23
 CC genes and open reading frames contained in an 80kb DNA sequence
 CC AAZ21501. Spinosyns are insecticidal microtides which are useful for the
 CC control of arachnids, nematodes and insects. Biosynthesis of spinosyns
 CC occurs via stepwise condensation and modification of carboxylic acid
 CC precursors generating a linear polyketide which is modified further. The
 CC DNA sequence contains a central region of approximately 55kb which has

CC homology to the DNA encoding the polyketide synthase (PKS) of known
 CC macroide producers. The spinosyn PKS DNA region consists of 5 ORFs with
 CC acop codons at the end of acyl carrier protein (ACP) domains. Together
 CC the PKS polypeptides (AAV39297-Y39301) of which this sequence is one,
 CC form a complex consisting of an initiator module, spnA, and several
 CC extender modules spnB-spnF. Each extender module adds a specific acetyl
 CC Co-A precursor to a growing polyketide chain, and modifies the beta-keto
 CC group in a specific manner. A module in a PKS polypeptide consists of
 CC several domains with specific functions. The initiator module has an
 CC acyl transferase (AT) domain, and an acyl carrier protein (ACP) domain.
 CC The extender module has the same domains plus a beta-ketoreductase (KS)
 CC domain, and an enoyl reductase (ER) domain. The last extender module
 CC terminates with a thioester domain. The products of the genes present
 CC in the upstream region the PKS genes have been assigned names spnF-spnS
 CC AAV39302-Y39315 and are responsible for different modifications in
 CC spinosyn biosynthesis. There are also two ORFs ORF15 and ORF16 present
 CC immediately upstream of spnS, producing polypeptides AAV39316-Y39317,
 CC and two ORFs ORF1 and ORF2 present downstream of the PKS region
 CC producing polypeptides AAV39318-Y39319. The genes are useful to improve
 CC yield of spinosyns, and for creating new spinosyns e.g. by mutagenesis,
 CC or interruption of steps in spinosyn biosynthesis. The modified
 CC spinosyns may be a new insect control agent or serve as substrates for
 CC further chemical modification and the creation of new semi-synthetic
 CC spinosyns. The genes are also useful to isolate similar sequences from S.
 CC spinosa or other species by hybridization.

SO Sequence 2595 AA:

Query Match 47.3%; Score 765; DB 20; Length 2595;

Best Local Similarity 50.6%; Pred. No. 1,7e-60;
 Matches 157; Conservative 51; Mismatches 92; Indels 10; Gaps 3;

QY 1 VFVPPGQGWPMGRELDAADVFRESVRACCAAFAPYVDMSVEQVLRDSDPAPGLDRV 60
 |||||
 DB 527 VFVPPGQGWPMGRELDAADVFRESVRACCAAFAPYVDMSVEQVLRDSDPAPGLDRV 586
 QY 61 DVVQPTLFVAVMISLAALMRSQGVPCAVLGHSLGEIAAAVSGSLADARVVTLMSCA 120
 |||||
 DB 587 DVVQPTLFVAVMISLAALMRSQGVPCAVLGHSLGEIAAAVSGSLADARVVTLMSCA 646
 QY 121 QTTLAGTALVSAATPDELPRIPATWEDNPARLVAANGPRSTVSGAREAVADVA 180
 |||||
 DB 647 LTLAAGSGAMASLQHPAEVRQILLPMRD----RIGVAGVNGPSTLVSGDRAMABELLA 702
 QY 181 DLTAQVTRMTIPVDVPAHSPMLVAIEERVVSGLLPTTPRPSRIPPHSSVTGRLDTREL 240
 |||||
 DB 703 ECADRELRMRRIPVEVASHSPHIEVDELGLGLAPVEPRGSIPIYSTTTGDLDP-RPM 761
 QY 241 DAAVYRNMSSTVAFEPARLLLOQGPKTVEVSPHVLTMGLQELAPDGLDGTGTADTV 300
 |||||
 DB 762 DADYVTRNLKQPVLFDAVBAALVKRGYDAFIEISPHVLTANTIOETAVRAG-----REV 816
 QY 301 IMGTLLRRGQG 310
 :|||
 DB 817 ALGTLRRGEG 826

RESULT 13

AAV70965 standard; Protein; 2595 AA.

AAV70965;

28-AUG-2001 (first entry)

S. spinosa protein fragment encoded by ORF18, SEQ ID 41.

XX Forosamine; trimethylxanthose; polyketide synthase; biosynthesis;
 XX spinosyn; polyketide aglycone; transgenic plant; insect resistance;
 XX macroide; insecticidal; polyketide synthase.
 OS Saccharopolyspora spinosa.

XX DE19957268-A1.
 XX 08-MAR-2001.
 XX 29-NOV-1999; 99DE-1057268.
 XX 27-AUG-1999; 99DE-1040596.
 XX (FARB) BAYER AG.
 XX Ebertz G, Noehle V, Froede R, Velten R, Salas JA;
 XX WPI; 2001-267102/28.
 XX N-PSDB; AAF68335.
 PT New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
 PT recombinant production of insecticidal spinosyns and their derivatives
 PS Claim 57; Page 179-189; 354pp; German.

CC This invention describes a novel method nucleic acid (I) and its encoded
 CC polypeptide (II) containing at least one region that encodes an enzymatic
 CC activity involved in biosynthesis of spinosyns. (I) are used (1) to
 CC identify, inactivate or modulate genes involved in the biosynthesis of
 CC (II); (ii) to generate a library of polyketide synthases; (iii) for
 CC adding forosamine or trimethylxanthose to a spinosyn or polyketide
 CC aglycone; and (iv) for recombinant production of the corresponding
 CC enzymes, which are used for production of (II), their precursors or
 CC derivatives, including production of transgenic plants that express (II)
 CC and thus have increased resistance to insects. (I) are also useful as
 CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
 CC macroides with insecticidal, but not antibacterial, activity, and can
 CC also be used to raise specific antibodies, useful for identifying
 CC expression clones in a gene bank. Cells transformed with (I) may produce
 CC (II) at significantly increased levels or produce new derivatives of
 CC (II). This sequence represents an S. spinosa polyketide synthase.

SO Sequence 2595 AA:

Query Match 47.3%; Score 765; DB 22; Length 2595;

Best Local Similarity 50.6%; Pred. No. 1,7e-60;
 Matches 157; Conservative 51; Mismatches 92; Indels 10; Gaps 3;

QY 1 VFVPPGQGWPMGRELDAADVFRESVRACCAAFAPYVDMSVEQVLRDSDPAPGLDRV 60
 |||||
 DB 527 VFVPPGQGWPMGRELDAADVFRESVRACCAAFAPYVDMSVEQVLRDSDPAPGLDRV 586
 QY 61 DVVQPTLFVAVMISLAALMRSQGVPCAVLGHSLGEIAAAVSGSLADARVVTLMSCA 120
 |||||
 DB 587 DVVQPTLFVAVMISLAALMRSQGVPCAVLGHSLGEIAAAVSGSLADARVVTLMSCA 646
 QY 121 QTTLAGTALVSAATPDELPRIPATWEDNPARLVAANGPRSTVSGAREAVADVA 180
 |||||
 DB 647 LTLAAGSGAMASLQHPAEVRQILLPMRD----RIGVAGVNGPSTLVSGDRAMABELLA 702
 QY 181 DLTAQVTRMTIPVDVPAHSPMLVAIEERVVSGLLPTTPRPSRIPPHSSVTGRLDTREL 240
 |||||
 DB 703 ECADRELRMRRIPVEVASHSPHIEVDELGLGLAPVEPRGSIPIYSTTTGDLDP-RPM 761
 QY 241 DAAVYRNMSSTVAFEPARLLLOQGPKTVEVSPHVLTMGLQELAPDGLDGTGTADTV 300
 |||||
 DB 762 DADYVTRNLKQPVLFDAVBAALVKRGYDAFIEISPHVLTANTIOETAVRAG-----REV 816
 QY 301 IMGTLLRRGQG 310
 :|||
 DB 817 ALGTLRRGEG 826

RESULT 14

AAW22603 standard; Protein; 3729 AA.

FT	/label= AT11 domain	
FT	/note= "Acyltransferase (AT) domain"	
FT	4231..4432	
FT	/label= DH11 domain	
FT	/note= "Dehydratase (DH) domain"	
FT	4523..4770	
FT	/label= KR11 domain	
FT	/note= "Ketoreductase (KR) domain"	
FT	4806..4879	
FT	/label= ACP11 domain	
FT	/note= "Acy1 carrier protein (ACP) domain"	
FT	4801..5325	
FT	/label= KS12 domain	
FT	/note= "Ketosynthase (KS) domain"	
FT	5432..5754	
FT	/label= AT12 domain	
FT	/note= "Acy1transferase (AT) domain"	
FT	5768..5977	
FT	/label= DH12 domain	
FT	/note= "Dehydratase (DH) domain"	
FT	6068..6315	
FT	/label= KR12 domain	
FT	/note= "Ketoreductase (KR) domain"	
FT	6348..6421	
FT	/label= ACP12 domain	
FT	/note= "Acy1 carrier protein (ACP) domain"	
FT	6434..6873	
FT	/label= KS13 domain	
FT	/note= "Ketosynthase (KS) domain"	
PN	WO200159126-A2.	
XX	16-AUG-2001.	
XX	08-FEB-2001; 2001WO-GB00509.	
XX	08-FEB-2000; 2000GB-0002840.	
PR	10-APR-2000; 2000GB-0008786.	
PR	14-APR-2000; 2000GB-0009387.	
XX	(UYNO-) UNIV NORGES TEKNIISK NATURVITENSKAPELIGE.	
PA	(SNTF) SINTER STIPELSEN IND TEK FORSK.	
PA	(ALPH-) ALPHARMA AS.	
PA	(SINV-) SIVENT AS.	
PA	(DZIE/) DZIEGLEWSKA H.	
PA	(ZOTC/) ZOTCHEV S B.	
PA	(SEKU/) SEKUROVA O N.	
PA	(FJAE/) FJAEVYK E.	
PA	(BRAU/) BRAUTASSET T.	
PA	(STRO/) STROM A R.	
XX	Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;	
PI	Valla S, Ellingsen TE, Sietta H, Guilleken O;	
PI	WPI: 2001-557614/62.	
DR	N-Psdb; AAD17185.	
XX	New nystatin polyketide synthase polynucleotides and polypeptides,	
XX	useful as antibiotics and antifungals -	
PS	claim 15; Page 181-185; 266pp; English.	
XX	The present invention relates to the cloning and sequencing of the gene	
CC	cluster encoding a modular type I polyketide synthase (PKS) enzyme	
CC	involved in the biosynthesis of the macrolide antibiotic nystatin.	
CC	The nystatin PKS is useful as antifungal antibiotic. The present	
CC	sequence is a PKS type I encoding Streptomyces noursei	
CC	nystatin gene, NysI (partial).	
XX	Sequence 7068 AA;	
XX	10	

Matches	168;	Conservative	34;	Mismatches	105;	Indels	10;	Gaps	3;
Qy	1	VVFVPGGQPGWPKGREGELLDASDVFRSSVACEAALPAPYDWMSEYQVLRLDSDPAAGLDLV							60
Db	3895	VFEVPGGSGQWVGMSQILLDBESAFAFARIACEAALAEFTDMSLVLDVLRGVAGPSLERY							3954
Qy	61	DVVOPTLFVAVMISLAALMBRSQGVPCAVLGHSGEIAAAHVSGSLSLADARVVTWLSQA							120
Db	3955	DVVOPASFAVWVSLAALMBRSRGVLDPDVGHSQGEIAAAVSGSLSLRDGARVVALRSQA							4014
Qy	121	-QTTLAGTGAALVSAATPDELPRPIAEWTEQNPARTLVAANVNGRSTVSGAREAVADLV							179
Db	4015	IGRLAARGGMSVALSVDVLERLVLF- - - - -EGRVAAVANGRSVVVAGEPALDALH							4070
Qy	180	ADLTAAOVRTMTFVDVPAHSPMTVAIEERVVSGGLPIPTPPSHIPHSSVTGRLDTRE							239
Db	4071	ARLTADDIRARRIADVASHSHQVEDHEBELLEVLAEIAPRTSEVPFFSVYTGDMLDIAR							4130
Qy	240	LDAAYWRNMSSTVREFPARLLILQOQPKTFVESHSPHVLVTMGLOELAPDLGDTTGTADT							299
Db	4131	MDAAYWFRNLRGRGRREADAVALDLAAEYRAFEVESHSPVLSMAQOEAIDEAG- - - - -VPA							4185
Qy	300	VIMGTLRRGQGTLDHFL 316							
Db	4186	VAAGTILRRDQGTDRFL 4202							

Search completed: June 17, 2003, 13:02:27
Job time : 16.48 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 12:53:02 ; Search time 6.74779 Seconds
(without alignments)
4458.180 Million cell updates/sec

Title: US-09-914-286-3_COPY_2983_3128

Perfect score: 764
Sequence: 1 YMESTOPGAGNVSAAGLDP.....ACGGGASGGVWTCASGVL 146

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: SP archaea:*
- 2: SP bacteria:*
- 3: SP fungi:*
- 4: SP human:*
- 5: SP_invertebrate:*
- 6: SP_mammal:*
- 7: SP_mhc:*
- 8: SP_organelle:*
- 9: SP_phase:*
- 10: SP_plant:*
- 11: SP_virus:*
- 12: SP_virus:*
- 13: SP_vertebrate:*
- 14: SP_unclassified:*
- 15: SP_virus:*
- 16: SP_bacteriophage:*
- 17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	764	100.0	3972 2 Q9S0R8	Q9S0R8 streptomyc
2	760	99.5	3626 2 Q9F7T9	Q9F7T9 streptomyc
3	585.5	76.6	5532 2 Q9S0R4	Q9S0R4 streptomyc
4	387.5	50.7	3576 16 Q9EX53	Q9EX53 streptomyc
5	386.5	50.6	3524 2 Q93H86	Q93H86 streptomyc
6	386	50.5	10917 2 Q93NM6	Q93NM6 streptomyc
7	385.5	50.5	3739 2 Q9ZG14	Q9ZG14 streptomyc
8	381	49.9	9510 2 Q93NX9	Q93NX9 streptomyc
9	377	49.3	9507 2 Q9EW11	Q9EW11 streptomyc
10	372	48.7	9477 2 Q9L4X3	Q9L4X3 streptomyc
11	366.5	48.0	3970 2 Q93H18	Q93H18 streptomyc
12	365.5	47.8	5588 2 Q9ALM2	Q9ALM2 saccharopol
13	362	47.4	6396 2 Q9KID7	Q9KID7 streptomyc
14	361.5	47.3	3591 2 Q9K1E1	Q9K1E1 streptomyc
15	361	47.3	6420 2 P95814	P95814 streptomyc
16	357	46.7	4881 2 Q9S0R3	Q9S0R3 streptomyc

17	357	46.7	7525 2 Q9K1E0	Q9K1E0 streptomyc
18	354.5	46.4	11096 2 Q9L4W3	Q9L4W3 streptomyc
19	351.5	46.0	6239 2 Q9S0R7	Q9S0R7 streptomyc
20	351	45.9	3939 2 Q93HJ3	Q93HJ3 streptomyc
21	350	45.8	6048 2 Q93H87	Q93H87 streptomyc
22	348	45.5	6797 2 Q9X393	Q9X393 streptomyc
23	346.5	45.4	2103 16 Q9CD81	Q9CD81 mycobacteri
24	342	44.8	2152 2 Q9ALM5	Q9ALM5 saccharopol
25	337.5	44.2	1907 2 Q93H19	Q93H19 streptomyc
26	336.5	44.0	6315 2 Q9ADL6	Q9ADL6 polyangium
27	333.5	43.7	3613 2 Q93HJ1	Q93HJ1 streptomyc
28	333.5	43.7	4472 2 Q93H54	Q93H54 streptomyc
29	332.5	43.5	1808 2 Q9EWA2	Q9EWA2 streptomyc
30	332	43.5	5069 2 Q52789	Q52789 amycolatops
31	329	43.1	5062 2 Q52545	Q52545 amycolatops
32	329	43.1	10223 2 Q54296	Q54296 streptomyc
33	324.5	42.5	4151 16 Q53490	Q53490 mycobacteri
34	322.5	42.2	2024 2 Q9EWA3	Q9EWA3 streptomyc
35	320	41.9	4613 2 Q9ZG15	Q9ZG15 streptomyc
36	319	41.8	3816 2 Q9K1V3	Q9K1V3 streptomyc
37	317	41.5	7576 2 Q9ZG44	Q9ZG44 streptomyc
38	315.5	41.3	1937 2 Q30482	Q30482 streptomyc
39	314.5	41.2	8653 2 Q54297	Q54297 streptomyc
40	313.5	41.0	8817 2 Q53840	Q53840 polyangium
41	312	40.8	2126 16 P94996	P94996 mycobacteri
42	311.5	40.8	4928 2 Q9ALM3	Q9ALM3 saccharopol
43	311.5	40.8	5644 2 Q93NX8	Q93NX8 streptomyc
44	309.5	40.5	5435 2 Q9L4X2	Q9L4X2 streptomyc
45	308.5	40.4	3413 2 Q54593	Q54593 amycolatops

ALIGNMENTS

RESULT 1

Q9S0R8 PRELIMINARY; PRT; 3972 AA.
ID Q9S0R8
AC Q9S0R8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 1.
GN AVEA1.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9380548; PubMed=10449723;
RA Ikeda H., Nonomiyama T., Usami M., Ohta T., Omura S.;
RT "Organization of the biosynthetic gene cluster for the polyketide
antihelminthic macrocyclic avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL; AB032367; BAA84474.1; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantane-attach.
DR InterPro; IPR000834; Zn_carboxypept.
DR Pfam; PF00698; Acyl_transf. 3.
DR Pfam; PF00109; ketoacyl-synt. 2.
DR Pfam; PF02801; ketoacyl-synt. C' 2.
DR Pfam; PF00550; pp-binding; 3.
DR PROSITE; PS50075; ACP DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL SYNTHASE; 2.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_2.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 3972 AA; 416852 MW; 2A2936958032B1C3 CRC64;

Query Match 100.0%; Score 764; DB 2; Length 3972;
Best Local Similarity 100.0%; Pred. No. 9.2e-51;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMLESTPGAGNVSAAGLDPTHEHLGATLELATDGGALLAGRLSLRSHPLADHAAGT 60
 DB 2983 YMLESTPGAGNVSAAGLDPTHEHLGATLELATDGGALLAGRLSLRSHPLADHAAGT 3042
 QY 61 VLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGVSVQVGVAAADGEGRLVSV 120
 DB 3043 VLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGVSVQVGVAAADGEGRLVSV 3102
 QY 121 YARGSAACGGGASGGVWTCASGVL 146
 DB 3103 YARGSAACGGGASGGVWTCASGVL 3128

RESULT 2

Q9F779 PRELIMINARY; PRT; 3626 AA.
 AC Q9F779;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Avermectin polyketide synthase (fragment).
 OS Streptomyces avermectilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
 NX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC31271;
 RA Hong Y.-S., Lee J.-J.;
 RT "Targeted Gene Disruption of the avermectin O-methyltransferase gene and polyketide synthase gene from Streptomyces avermectilis";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF259943; AAC09812.1; -
 DR InterPro: IPR001227; Ac_transferase.
 DR InterPro: IPR00794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantene_attach.
 DR Pfam: PF00698; Acyl_transf. 3.
 DR Pfam: PF00109; Ketoacyl-synt. 2.
 DR Pfam: PF02801; Ketoacyl-synt. C; 2.
 DR Pfam: PF00550; PP-binding; 2.
 DR PROSITE: PS50075; ACP_DOMAIN; 2.
 DR PROSITE: PS00066; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 FT NON TER 3626
 SQ SEQUENCE 3626 AA; 380557 MW; 6272F5F088C1A8D0 CRC64;

Query Match 99.5%; Score 760; DB 2; Length 3626;
 Best Local Similarity 99.3%; Pred. No. 1.7e-50;
 Matches 145; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YMLESTPGAGNVSAAGLDPTHEHLGATLELATDGGALLAGRLSLRSHPLADHAAGT 60
 DB 2950 YMLESTPGAGNVSAAGLDPTHEHLGATLELATDGGALLAGRLSLRSHPLADHAAGT 3009
 QY 61 VLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGVSVQVGVAAADGEGRLVSV 120
 DB 3010 VLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGVSVQVGVAAADGEGRLVSV 3069
 QY 121 YARGSAACGGGASGGVWTCASGVL 146
 DB 3070 YARGSAACGGGASGGVWTCASGVL 3095

RESULT 3

Q9S0R4 PRELIMINARY; PRT; 5532 AA.
 AC Q9S0R4;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAY-2002 (Tremblrel. 20, Last annotation update)
 DE Type I polyketide synthase AVES 3.
 GN AVEA3.

OS Streptomyces avermectilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
 NX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=99380548; PubMed=10449723;
 RA Ikeda H., Nonomura T., Usami M., Ohta T., Omura S.;
 RT "Organization of the biosynthetic gene cluster for the polyketide antihelminthic macroide avermectin in Streptomyces avermectilis";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
 DR EMBL: AB032367; BA04478.1; -
 DR InterPro: IPR001227; Ac_transferase.
 DR InterPro: IPR00794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantene_attach.
 DR InterPro: IPR000834; Zn_carbOpept.
 DR Pfam: PF00698; Acyl_transf. 3.
 DR Pfam: PF00109; Ketoacyl-synt. 3.
 DR Pfam: PF02801; Ketoacyl-synt. C; 3.
 DR Pfam: PF00550; PP-binding; 3.
 DR PROSITE: PS50075; ACP_DOMAIN; 3.
 DR PROSITE: PS00066; B_KETOACYL_SYNTHASE; 3.
 DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 3.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 5532 AA; 575193 MW; EC7B2069D44A057 CRC64;

Query Match 76.6%; Score 585.5; DB 2; Length 5532;
 Best Local Similarity 76.0%; Pred. No. 1e-36;
 Matches 111; Conservative 15; Mismatches 19; Indels 1; Gaps 1;

QY 1 YMLESTPGAGNVSAAGLDPTHEHLGATLELATDGGALLAGRLSLRSHPLADHAAGT 60
 DB 955 YMLDA-PTGAGDVAAAGLDEAHPHPLAATVQGLDTGCLTGLSLRSHWLDYEVGGA 1013
 QY 61 VLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGVSVQVGVAAADGEGRLVSV 120
 DB 1014 VLLSGAFVELALAVGSEVGCRTIEQITVHAPLVVPDGVSVQVGVAAADGEGRLVSV 1073
 QY 121 YARGSAACGGGASGGVWTCASGVL 146
 DB 1074 YARGSAACGGGASGGVWTCASGVL 1099

RESULT 4

Q9EX53 PRELIMINARY; PRT; 3576 AA.
 AC Q9EX53;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Putative type I polyketide synthase.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
 NX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kleiser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;

"A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN: A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornaby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Welterorek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces

coelicolor A3(2)." (2002).

RL Nucleic Acids Res. 30:1471-1477(2002).

DR EMBL/AL512902; CAC22145.1; -.

DR HSSP; P25715; 1MLA.

DR InterPro: IPR004410; Pfam.

DR InterPro: IPR000794; Ketoacyl-synt.

DR InterPro: IPR003880; Ppantne_attach.

DR InterPro: IPR003662; sub_transporter.

DR InterPro: IPR001998; Xylose_isom.

DR Pfam; PF00698; Acyl_transf_2.

DR Pfam; PF02801; ketoacyl-synt_C; 2.

DR Pfam; PF00550; pp-binding; 2.

DR TIGRfam; TIGR00128; fadD; 2.

DR PROSITE; PS50075; ACP_DOMAIN; 2.

DR PROSITE; PS00012; B_KETOACYL_SYNTHASE; 2.

DR PROSITE; PS00215; SUGAR_TRANSPORT_1; UNKNOWN_2.

DR PROSITE; PS00173; XYLOSE_ISOMERASE_2; UNKNOWN_1.

KW Phosphopantetheine.

SEQUENCE 3576 AA; 377350 MW; 8A83C6F90AC9928B CRC64;

Query Match 50.7%; Score 387.5; DB 16; Length 3576;

Best Local Similarity 54.7%; Pred. No. 1.7e-21;

Matches 81; Conservative 15; Mismatches 49; Indels 3; Gaps 2;

QY 1 YMLESTPGAGNVSAAGLDPTBHPGLGATLELATDGCALLAGRLSLRSHPLADHAYGVT 60

DB 921 YMLDVTB-AAADAAGLGLTRDHPILGATIDLAGCEGQLTFSRLSTKTHPLAHHTYAGT 979

QY 61 VLSGATFLEIALHAGTYVGCGRVDELTLHAPLVVPVDSGVSVQVGAADGEGRLVSV 120

DB 980 TLLPGTGFVELAALAGRLGCPREBELTSLAPVLPERDGVRIQLAVGEADGAGRIVDV 1039

QY 121 YAR--GGSACGGGASGAGVWTCASGVL 146

DB 1040 YARPDDGETPAGAVAKARPMATLAKGVL 1067

RESULT 5

Q93H86 PRELIMINARY; PRT; 3524 AA.

AC Q93H86; 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Modular polyketide synthase.

OS Streptomyces avermitilis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomyces; Streptomyces; Streptomyces.

OX NCBI_TaxID=33903;

RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: Deducing the ability of producing secondary
 RT metabolites.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

DR EMBL/AB070949; BAB69304.1; -.

DR InterPro: IPR001227; Ac transferase.

DR InterPro: IPR000794; Ketoacyl-synt.

DR InterPro: IPR001993; Mitoch_carrier.

DR InterPro: IPR003880; Ppantne_attach.

DR Pfam; PF00698; Acyl_transf_2.

DR Pfam; PF02801; ketoacyl-synt_C; 2.

DR Pfam; PF00550; pp-binding; 2.

DR PROSITE; PS50075; ACP_DOMAIN; 2.

DR PROSITE; PS00012; B_KETOACYL_SYNTHASE; UNKNOWN_2.

DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.

KW Phosphopantetheine.

SEQUENCE 3524 AA; 36619 MW; 64A65759C461EBC4 CRC64;

Query Match 50.6%; Score 386.5; DB 2; Length 3524;

Best Local Similarity 56.2%; Pred. No. 1.9e-21;

Matches 82; Conservative 16; Mismatches 39; Indels 9; Gaps 2;

QY 1 YMLESTPGAGNVSAAGLDPTBHPGLGATLELATDGCALLAGRLSLRSHPLADHAYGVT 60

DB 2611 YMLEBAGAP-AGDRAAGLADHPFJSAAYELPDAGSHLTLGLSLQAPWMLADHAYGVT 2669

QY 61 VLSGATFLEIALHAGTYVGCGRVDELTLHAPLVVPVDSGVSVQVGAADGEGRLVSV 120

DB 2670 VLLPGTGFVELAALAGRLGCPREBELTSLAPVLPERDGVRIQLAVGEADGAGRIVDV 2729

QY 121 YARGGSAACGGGASGAGVWTCASGVL 146

DB 2730 HSRAG-----DDGPWIRNATGVL 2747

RESULT 6

Q93H86 PRELIMINARY; PRT; 10917 AA.

AC Q93H86; 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE AmpHC.

OS Streptomyces nodosus.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomyces; Streptomyces; Streptomyces.

OX NCBI_TaxID=40318;

RP SEQUENCE FROM N.A.
 RA Caffrey P., Lynch S.V., Flood E.M., Finnan S.M., O'Leary M.;
 RT "The amphoterich biosynthetic gene cluster from Streptomyces
 RT nodosus.";

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL/AE57202; AAK73514.1; -.

DR InterPro: IPR001227; Ac transferase.

DR InterPro: IPR002328; ADH_zinc.

DR InterPro: IPR002085; Adh_zn_family.

DR InterPro: IPR004410; fadD.

DR InterPro: IPR000794; ketoacyl-synt.

DR InterPro: IPR003880; Ppantne_attach.

DR Pfam; PF00698; Acyl_transf_6.

DR Pfam; PF00107; adh_zinc; 1.

DR Pfam; PF00109; ketoacyl-synt_C; 6.

DR Pfam; PF02801; ketoacyl-synt_C; 6.

DR Pfam; PF00550; pp-binding; 6.

DR TIGRfam; TIGR00128; fadD; 6.

DR PROSITE; PS50075; ACP_DOMAIN; 6.

DR PROSITE; PS00059; ADH_ZINC; UNKNOWN_1.

DR PROSITE; PS00606; B KETOACYL SYNTHASE; UNKNOWN_6.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_5.
 KM Phosphopantetheine.
 SQ SEQUENCE 10917 AA; 1132905 MW; 15AC5956B5810A1 CRC64;

Query Match 50.5%; Score 386; DB 2; Length 10917;
 Best Local Similarity 56.2%; Pred. No. 7,3e-21;
 Matches 82; Conservative 10; Mismatches 48; Indels 6; Gaps 1;

QY 1 YWLESTPGAGNVAAGIDPTEHPLLGATLELATDGGALLAGRLSLRSHPLADHAVGCT 60
 DB 9958 FWPGRGAAGDVPAAAGLSAHEPLLGAAVELAABEGVLFTRGLSLSSHPLADHAVWGCT 10017
 QY 61 VLLSGATFLELALHAGTYGCDRVDELTLHAPLVVPDGGVSVGVAAADGEGRLVSV 120
 DB 10018 VLLPGLTLLLENAVPAAGVEGCGRVEELTLAAPLVPERGALQVAVAAAPDGGRRIVGV 10077
 QY 121 YARGSGACGGGASGGVWTCASGVL 146
 DB 10078 YSR-----LEGADSVWAQHAAGVL 10097

RESULT 7
 Q9ZG14 PRELIMINARY; PRT; 3739 AA.
 AC Q9ZG14;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Type I polyketide synthase PkAII.
 GN PkAII.
 OS Streptomyces venezuelae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=54571;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC15439;
 RX MEDLINE=98445333; PubMed=9770448;
 RA Xue Y., Zhao L., Liu H.W., Sherman D.H.;
 RT "A gene cluster for macroide antibiotic biosynthesis in streptomyces
 venezuelae: architecture of metabolic diversity."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12111-12116(1998).
 DR EMBL: AF079138; AAC69330.1; -;
 DR HSP: O02054; 2AF8
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR002085; Ach_zn_family.
 DR InterPro: IPR00794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantane attach.
 DR InterPro: IPR002364; OOR_zeta_crystal.
 DR InterPro: IPR000169; SHPOT_acsite.
 DR Pfam: PF00698; Acyl_transf_2.
 DR Pfam: PF00107; adh_zinc_1.
 DR Pfam: PF00109; ketoacyl-synt_2.
 DR Pfam: PF02801; ketoacyl-synt_C_2.
 DR Pfam: PF00550; pp-binding_2.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 DR PROSITE; PS00606; B_KETOACYL SYNTHASE; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
 DR PROSITE; PS01162; OOR_ZETA_CRYSTAL; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 3739 AA; 387170 MW; 3D3910824DA5B080 CRC64;

Query Match 50.5%; Score 385.5; DB 2; Length 3739;
 Best Local Similarity 54.8%; Pred. No. 2.5e-21;
 Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YWLESTPGAGNVAAGIDPTEHPLLGATLELATDGGALLAGRLSLRSHPLADHAVGCT 60
 DB 2406 YWPPDLSAAGDITSAAGLAHPLLGAAVALADSDCLTGLSISTRTHEPLADHAVAGT 2465
 QY 61 VLLSGATFLELALHAGTYGCDRVDELTLHAPLVVPDGGVSVGVAAADGEGRLVSV 120

DB 2466 VLLPGLTFAVELARAGDQVCDLVEELTLDAPLVLPBRGAVRQLSVGASDESGRRVFG 2525
 QY 121 YARGSGACGGGASGGVWTCASGVL 146
 DB 2526 YAHPEADAPG-----EAETWTHATGVL 2546

RESULT 8
 Q93NX9 PRELIMINARY; PRT; 9510 AA.
 ID Q93NX9;
 AC Q93NX9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE AmpH1.
 GN AMPH1.
 OS Streptomyces nodosus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=40318;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Catfrey P., Lynch S.V., Flood E.M., Finnan S.M., O'Leary M.;
 RT "The amphoterich biosynthetic gene cluster from Streptomyces
 nodosus."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF357202; AK73501.1; -;
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR004410; FAD.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR001005; Myd_DNA_binding.
 DR InterPro: IPR003880; Pantane attach.
 DR Pfam: PF00698; Acyl_transf_1.
 DR Pfam: PF00106; adh_shortf_1.
 DR Pfam: PF00109; ketoacyl-synt_C_6.
 DR Pfam: PF02801; ketoacyl-synt_C_6.
 DR Pfam: PF00550; pp-binding_6.
 DR TIGRFAMs: TIGR00128; fadD; 6.
 DR PROSITE; PS50075; ACP DOMAIN; 6.
 DR PROSITE; PS00606; B_KETOACYL SYNTHASE; UNKNOWN_6.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE; PS00037; MYR_1; UNKNOWN_1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_6.
 KM Phosphopantetheine.
 SQ SEQUENCE 9510 AA; 992817 MW; 1A8E003A15D478CF CRC64;

Query Match 49.9%; Score 381; DB 2; Length 9510;
 Best Local Similarity 53.4%; Pred. No. 1.5e-20;
 Matches 78; Conservative 13; Mismatches 49; Indels 6; Gaps 1;

QY 1 YWLESTPGAGNVAAGIDPTEHPLLGATLELATDGGALLAGRLSLRSHPLADHAVGCT 60
 DB 907 YWPRVAGGADLRSGVGLAAHPLLSAAVSLADSSGALLTGRISLSHPLADHAVTGA 966
 QY 61 VLLSGATFLELALHAGTYGCDRVDELTLHAPLVVPDGGVSVGVAAADGEGRLVSV 120
 DB 967 TLLPGLTFAVELARAGDQVCDRVDELTLAAPLVLEGGGVQVQVLMIGNPDAGRRSVTV 1026
 QY 121 YARGSGACGGGASGGVWTCASGVL 146
 DB 1027 YGRPDA-----DEDAFWTSHATGVL 1046

RESULT 9
 Q9EWAL PRELIMINARY; PRT; 9507 AA.
 ID Q9EWAL;
 AC Q9EWAL;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE P1ms2 protein.
 GN P1ms2.
 OS Streptomyces natalensis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=68242;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20547809; PubMed=11094342;
 RT "A complex multienzyme system encoded by five polypeptide synthase genes is involved in the biosynthesis of the 26-membered polyene macroide pimaricin in Streptomyces natalensis."
 RL Chem. Biol. 7:895-905(2000).
 DR EMBL; AJ278573; CAC20921.1; -.
 DR HSSP; P25715; IMLA.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR004410; Fabd.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR001005; Myb_DNA_binding.
 DR InterPro; IPR003880; Ppantne_attach.
 DR Pfam; PF00698; Acyl_transfer; 6.
 DR Pfam; PF00109; ketoacyl-synt; 6.
 DR Pfam; PF02801; ketoacyl-synt; C; 6.
 DR Pfam; PF00550; pp-binding; 6.
 DR TIGRfam; TIGR00128; fabd; 6.
 DR PROSITE; PS50075; ACP_DOMAIN; 6.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 6.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 5.
 KM Phosphopantetheine.
 SQ SEQUENCE 9507 AA; 994228 MW; 57BD80C8AD3789F CRC64;
 Query Match 49.3%; Score 377; DB 2; Length 9507;
 Best Local Similarity 54.8%; Pred. No. 3,2e-20;
 Matches 80; Conservative 13; Mismatches 47; Indels 6; Gaps 1;
 QY 1 YWLESTPGAGNVSAAGLDPTFPLGATLELATDGGALLAGRLSRSHPMADHAYG 60
 DB 913 YMPKSLPATGGDVAAGAGAAHHPPLTAASVANSDDGLLTGRLSRTHPLADHAYRG 972
 QY 61 VLSGATFLELALHAGTYGCDRVDLTLPALVPVVDGVSVOVGAAADGGRRLVSV 120
 DB 973 VLLPGTAFLELAVAGDEAGGRVLELTLPALVLPBEGGVQVQVWVGSPDASGRAVS 1032
 QY 121 YARGSGAGGAGSGGVWTCASGVL 146
 DB 1033 HSR-----PDGPPELPTQHAAGTL 1052
 RESULT 10
 Q914X3 PRELIMINARY; PRT; 9477 AA.
 ID Q914X3;
 AC Q914X3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE NYSI.
 GN NYSI.
 OS Streptomyces nourset.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1971;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=ATCC 11455;
 RX MEDLINE=20334850; PubMed=10873841;
 RA Brautaset T., Sekurova O.N., Sletta H., Ellingsen T.E., Strom A.R.,
 Valla S., Zotchev S.B.,

RT "Biosynthesis of the polyene antifungal antibiotic nystatin in Streptomyces nourset ATCC 11455: analysis of the gene cluster and RT deduction of the biosynthetic pathway."
 RL Chem. Biol. 7:395-403(2000).
 DR EMBL; AF263912; AF711766.1; -.
 DR HSSP; P81989; 10E5.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR004410; Fabd.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR001005; Myb_DNA_binding.
 DR InterPro; IPR003880; Ppantne_attach.
 DR Pfam; PF00698; Acyl_transfer; 6.
 DR Pfam; PF00109; ketoacyl-synt; 3.
 DR Pfam; PF00109; ketoacyl-synt; 6.
 DR Pfam; PF02801; ketoacyl-synt; C; 6.
 DR Pfam; PF00550; pp-binding; 6.
 DR TIGRfam; TIGR00128; fabd; 6.
 DR PROSITE; PS50075; ACP_DOMAIN; 6.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 6.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 6.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 9477 AA; 988132 MW; 0D2BCA5D7B265483 CRC64;
 Query Match 48.7%; Score 372; DB 2; Length 9477;
 Best Local Similarity 51.4%; Pred. No. 7.7e-20;
 Matches 75; Conservative 18; Mismatches 47; Indels 6; Gaps 1;
 QY 1 YWLESTPGAGNVSAAGLDPTFPLGATLELATDGGALLAGRLSRSHPMADHAYG 60
 DB 914 YWKAQSGTADLRSGVGLGAHHPPLSAVSLADAGSTLTGRLSRTHPLADHAYRG 973
 QY 61 VLSGATFLELALHAGTYGCDRVDLTLPALVPVVDGVSVOVGAAADGGRRLVSV 120
 DB 974 TLLPGTAFLELAVAGDEAGGRVLELTLPALVLPBEGGVQVQVWVGSPDASGRAVS 1033
 QY 121 YARGSGAGGAGSGGVWTCASGVL 146
 DB 1034 HARPD-----GDDTPTWTAHATGVL 1053
 RESULT 11
 Q93H18 PRELIMINARY; PRT; 3970 AA.
 ID Q93H18;
 AC Q93H18;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Modular polyketide synthase.
 GN OLM44.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Iehikawa J., Hanamoto A., Takahashi C.,
 RA Shinoe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
 RT "Genome sequence of an industrial microorganism Streptomyces RT avermitilis: Deducing the ability of producing secondary RT metabolites."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL; AB070940; BAB69199.1; -.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh_zn_family.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Ppantne_attach.
 DR InterPro; IPR002364; QOR_zeta_crystal.

DR Pfam; PF00698; Acyl transferase; 2.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketocacyl-synt; 2.
 DR Pfam; PF02801; ketocacyl-synt_C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PS50075; ACP_DOMAIN; 2.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 DR PROSITE; PS01162; COO_ZETA_CRYSTAL; UNKNOWN_1.
 KW Phosphopantetheine.
 SQ SEQUENCE 3970 AA; 412657 MW; A93B583FAFA68C47 CRC64;

Query Match 48.0%; Score 366.5; DB 2; Length 3970;
 Best Local Similarity 53.4%; Pred. No. 8e-20;
 Matches 78; Conservative 14; Mismatches 47; Indels 7; Gaps 2;

QY 1 YWLESTQPGAGNVAAGLDPTHEPLLGATLELATDGGALLAGRLSLRSHPMADHAVGCT 60
 ID 09ALM2 PRELIMINARY; PRT; 5588 AA.
 AC 09ALM2;
 DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DE Polypeptide synthase extender modules 8-10.
 GN SPNE.
 OS Saccharopolyspora spinosa.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;
 OC Saccharopolyspora;
 NC NCB1_TaxID=60894;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21257765; PubMed=11358695;
 RA Waldron C., Matsushima P., Rostek P.R. Jr., Broughton M.C.,
 RA Turner J., Madduri K., Crawford K.P., Merlo D.J., Baltz R.H.;
 RT "Cloning and analysis of the spinosad biosynthetic cluster of
 RL Saccharopolyspora spinosa.";
 RL Chem. Biol. 8:487-499(2001).
 DR EMBL; AY007564; AAG3262.1; -
 DR HSP; P25715; IMLA.
 DR InterPro; IPR001227; Ac-transferase.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR00794; Ketocacyl-synt.
 DR InterPro; IPR003880; Pantne attach.
 DR InterPro; IPR001031; Phosphotransferase.
 DR Pfam; PF00698; Acyl transferase; 3.
 DR Pfam; PF00109; ketocacyl-synt; 3.
 DR Pfam; PF02801; ketocacyl-synt_C; 3.
 DR Pfam; PF00550; pp-binding; 3.
 DR Pfam; PF00975; Thioesterase; 1.
 DR PROSITE; PS50075; ACP_DOMAIN; 3.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
 KW Phosphopantetheine; Transferase.
 SQ SEQUENCE 5588 AA; 585629 MW; C47F38F44BAC6A11 CRC64;

Query Match 47.8%; Score 365.5; DB 2; Length 5588;
 Best Local Similarity 47.9%; Pred. No. 1.4e-19;
 Matches 70; Conservative 21; Mismatches 50; Indels 5; Gaps 1;

QY 1 YWLESTQPGAGNVAAGLDPTHEPLLGATLELATDGGALLAGRLSLRSHPMADHAVGCT 60
 ID 09KID7 PRELIMINARY; PRT; 6396 AA.
 AC 09KID7;
 DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
 DE Polypeptide extender units.";
 GN FKBA.
 OS Streptomyces hygroscopicus var. ascomyceticus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
 NC NCB1_TaxID=132248;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20332220; PubMed=10863099;
 RA Wu K., Chung L., Revill W.P., Katz L., Reeves C.D.;
 RA "The FK520 gene cluster of streptomyces hygroscopicus var.
 RT ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual
 RL polyketide extender units.";
 RL Gene 251:81-90(2000).
 DR EMBL; AF235504; AAF86396.1; -
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh_zn family.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR00794; Ketocacyl-synt.
 DR InterPro; IPR003880; Pantne attach.
 DR InterPro; IPR00634; S/T_dehydrtase.
 DR InterPro; IPR003662; sub_transporter.
 DR Pfam; PF00698; Acyl transferase; 4.
 DR Pfam; PF00107; adh_zinc; 2.
 DR Pfam; PF00109; ketocacyl-synt; 4.
 DR Pfam; PF02801; ketocacyl-synt_C; 4.
 DR Pfam; PF00550; pp-binding; 4.
 DR PROSITE; PS50075; ACP_DOMAIN; 4.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 4.
 DR PROSITE; PS00165; DEHYDRATASE_SER_THR; UNKNOWN_1.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_3.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 KW Phosphopantetheine; Transferase.
 SQ SEQUENCE 6396 AA; 669864 MW; FF13BC2431D2E805 CRC64;

Query Match 47.4%; Score 362; DB 2; Length 6396;
 Best Local Similarity 51.4%; Pred. No. 3e-19;
 Matches 75; Conservative 15; Mismatches 46; Indels 10; Gaps 2;

QY 1 YWLESTQPGAGNVAAGLDPTHEPLLGATLELATDGGALLAGRLSLRSHPMADHAVGCT 60
 ID 09KID7 PRELIMINARY; PRT; 6396 AA.
 AC 09KID7;
 DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
 DE Polypeptide extender units.";
 GN FKBA.
 OS Streptomyces hygroscopicus var. ascomyceticus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
 NC NCB1_TaxID=132248;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20332220; PubMed=10863099;
 RA Wu K., Chung L., Revill W.P., Katz L., Reeves C.D.;
 RA "The FK520 gene cluster of streptomyces hygroscopicus var.
 RT ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual
 RL polyketide extender units.";
 RL Gene 251:81-90(2000).
 DR EMBL; AF235504; AAF86396.1; -
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh_zn family.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR00794; Ketocacyl-synt.
 DR InterPro; IPR003880; Pantne attach.
 DR InterPro; IPR00634; S/T_dehydrtase.
 DR InterPro; IPR003662; sub_transporter.
 DR Pfam; PF00698; Acyl transferase; 4.
 DR Pfam; PF00107; adh_zinc; 2.
 DR Pfam; PF00109; ketocacyl-synt; 4.
 DR Pfam; PF02801; ketocacyl-synt_C; 4.
 DR Pfam; PF00550; pp-binding; 4.
 DR PROSITE; PS50075; ACP_DOMAIN; 4.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 4.
 DR PROSITE; PS00165; DEHYDRATASE_SER_THR; UNKNOWN_1.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_3.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 KW Phosphopantetheine; Transferase.
 SQ SEQUENCE 6396 AA; 669864 MW; FF13BC2431D2E805 CRC64;

```

RESULT 14
O9KIE1 ID O9KIE1 PRELIMINARY; PRT; 3591 AA.
AC O9KIE1
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE PKBC.
GN PKBC.
OS Streptomyces hygroscopicus var. ascomyceticus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=312248;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20323220; Pubmed=10863099;
RA Wu K., Chung L., Revill W.P., Katz L., Reeves C.D.;
RT "The FK520 gene cluster of streptomyces hygroscopicus var.
RT ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual
RT polycetide extender units."
RL Gene 251:81-90(2000).
DR EMBL; AF235504; AAF63392.1; -.
DR HSSP; Q02054; 2AF8.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Pantane_attach.
DR InterPro; IPR003662; Sub_transporter.
DR Pfam; PF00698; Acyl_transf; 2.
DR Pfam; PF00107; adh_zinc; 1.
DR Pfam; PF00109; ketoacyl-synt; 2.
DR Pfam; PF02801; ketoacyl-synt_C; 2.
DR Pfam; PF00550; pp-binding; 2.
DR PROSITE; PS50075; ACP_DOMAIN; 2.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 3591 AA; 373910 MW; 6CBF2D8C9CBB3B0 CRC64;

Query Match 47.3%; Score 361.5; DB 2; Length 3591;
Best Local Similarity 53.7%; Pred. No. 1.8e-19;
Matches 79; Conservative 13; Mismatches 44; Indels 11; Gaps 3;

QY 1 YMLESTPGAGNVSAGLDPTHEPLGATLEL-ATDGGALLAGRLSLRSHPMADHVG 59
DB 2465 YMLEAA--GATDLSAAGLTGAHMPMLAAVTALPADDGGVLTGRISLRTHPMADHVRG 2522
QY 60 TVLISGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGVSVQVGVAAADGEGRLVS 119
DB 2523 TVLLPGTAFVFLVIRADDEVCGLVDELIVIESPLVVPVTAADVSVVEGADENGRPVT 2582
QY 120 YVARGSGACGGGASGGVMTCHASGV 146
DB 2583 VHARTEGT-----GSMTRHAGTGL 2601

RESULT 15
P95814 ID P95814 PRELIMINARY; PRT; 6420 AA.
AC P95814;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-JUN-1997 (Tremblrel. 04, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE FK506 polycetide synthase.
GN FK5A.
OS Streptomyces sp.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=19311;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=MA6548;
RX MEDLINE=97217427; Pubmed=9063448;
RA Mccamelli H., Cai S.J., Shafiee S.J., Billecon K.O.;
RT "Structural organization of a multifunctional polycetide synthase
RT involved in the biosynthesis of the macroide immunosuppressant
RT FK506."
RL Eur. J. Biochem. 244:74-80(1997).
RX EMBL; Y10438; CAAT1463.1; -.
DR HSSP; P25715; 1MLA.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Pantane_attach.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00698; Acyl_transf; 4.
DR Pfam; PF00107; adh_zinc; 2.
DR Pfam; PF00109; ketoacyl-synt; 4.
DR Pfam; PF02801; ketoacyl-synt_C; 4.
DR Pfam; PF00550; pp-binding; 4.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50075; ACP_DOMAIN; 4.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 4.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_3.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_3.
SQ SEQUENCE 6420 AA; 672222 MW; 25BCD2069456BEF1 CRC64;

Query Match 47.3%; Score 361; DB 2; Length 6420;
Best Local Similarity 50.0%; Pred. No. 3.6e-19;
Matches 73; Conservative 18; Mismatches 45; Indels 10; Gaps 2;

QY 1 YMLESTPGAGNVSAGLDPTHEPLGATLEL-ATDGGALLAGRLSLRSHPMADHVG 60
DB 4309 YMLEAT--GAADITAGLDTHTAHPMLAATALEPTNCGTVLTGRISLRTHPMADHTI 4366
QY 61 TVLISGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGVSVQVGVAAADGEGRLVS 120
DB 4367 VLLPGTAFVFLVIRADDEVCDDTIDIVLETPLTLPVTAADVLTVEGADATGHRPVS 4426
QY 121 YVARGSGACGGGASGGVMTCHASGV 146
DB 4427 HAR-----PAGTDTWTHAGTGL 4444

Search completed: June 17, 2003, 13:07:57
Job time : 7.74779 secs

```

THIS PAGE BLANK (08/10)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:52:22 ; Search time 1.59218 Seconds
(without alignments)
3803.313 Million cell updates/sec

Title: US-09-914-286-3_COPY_2983_3128

Perfect score: 764
Sequence: 1 YMLESTOPGAGNVSAAGLDP.....ACGGGAGAGGVTCHASGVL 146

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	246.5	32.3	3567	1 ERY2_SACER	Q03132 saccharopol
2	96.5	12.6	2110	1 MCAS_MYCHO	Q02251 mycobacteri
3	85	11.1	957	1 Y278_MYCTU	P56877 mycobacteri
4	85	11.1	1582	1 YU30_RALSO	Q6XV02 ralslonia s
5	85	11.1	3591	1 PHAB_BORPE	P12255 bordetella
6	84.5	11.1	389	1 AZAB_PROHA	Q19054 procavia ca
7	84.5	11.1	1901	1 Y208_MYCTU	Q53553 mycobacteri
8	84	11.0	451	1 OCT6_RAT	P21967 rattus norv
9	83	10.9	449	1 OCT6_MOUSE	P21952 rattus muscu
10	81.5	10.7	324	1 FABH_RHOCA	P30796 rhodobacter
11	80.5	10.5	498	1 Y118_MYCTU	Q50615 mycobacteri
12	80.5	10.5	914	1 MA22_MYCTU	Q06799 mycobacteri
13	79	10.3	460	1 MB48_MYCTU	Q93348 mycobacteri
14	77.5	10.1	226	1 DEOC_MYCLE	Q9CB45 mycobacteri
15	77.5	10.1	358	1 YW05_MYCTU	Q10394 mycobacteri
16	77.5	10.1	559	1 HUTU_PSEAB	Q9HU83 pseudomonas
17	77.5	10.1	904	1 DP01_MYCTU	Q03155 mycobacteri
18	77.5	10.1	1286	1 AIDA_ECOTI	Q03155 mycobacteri
19	77.5	10.1	1876	1 PPSA_MYCTU	Q03155 mycobacteri
20	76.5	10.0	356	1 LEXD_RALSO	Q8XZ11 ralslonia s
21	76.5	10.0	1270	1 DDH9_HUMAN	Q08211 homo sapien
22	76	9.9	481	1 LORI_MOUSE	P18165 mus musculi
23	75.5	9.9	448	1 OCT6_HUMAN	Q03052 homo sapien
24	74	9.7	546	1 NUSM_CHURE	P08738 chlamydomon
25	74	9.7	778	1 Y034_MYCTU	P11933 mycobacteri
26	74	9.7	1084	1 HD44_HUMAN	P56524 homo sapien
27	73.5	9.6	289	1 TRPB_RHIME	Q32323 mycobacteri
28	73.5	9.6	406	1 TRPB_RHIME	Q32323 mycobacteri
29	73.5	9.6	543	1 VP91_MYCTU	Q50615 mycobacteri
30	73.5	9.6	825	1 SE5_FAT	Q53003 rattus norv
31	73.5	9.6	2035	1 HMP2_YERBN	P48633 yersinia en
32	73	9.6	911	1 PERT_BORBR	Q03033 bordetella
33	72.5	9.5	142	1 PA2E_MOUSE	Q9J13 mus musculi

34	72.5	9.5	151	1 DUT_YERPE	Q8ZJ5 yersinia pe
35	72.5	9.5	388	1 ARGJ_CORGL	Q59280 corynebacte
36	72.5	9.5	426	1 HUTI_BACHD	Q9B64 bacillus ha
37	72.5	9.5	447	1 ODP2_RHIME	Q9J93 rhizobium m
38	72	9.4	436	1 B10A_MYCLE	P45488 mycobacteri
39	72	9.4	473	1 Y133_MYCLE	P37351 mycobacteri
40	72	9.4	477	1 MA2_HUMAN	P56270 homo sapien
41	71.5	9.4	330	1 HX11_HUMAN	P31314 homo sapien
42	71.5	9.4	397	1 PRLA_LYSEN	P00778 lysobacter
43	71.5	9.4	530	1 HLY_HAI17	P29143 halophilic
44	71	9.3	335	1 TRX_MYCTU	P52214 mycobacteri
45	71	9.3	341	1 MRW_RHIME	Q9Z114 rhizobium m

ALIGNMENTS

```

RESULT 1
ERY2_SACER STANDARD; PRT; 3567 AA.
ID ERY2_SACER
AC Q03132; 054096;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Erythronolide synthase, modules 3 and 4 (EC 2.3.1.94) (ORF 2) (6-
DE deoxyerythronolide B synthase II) (DEBS 2).
GN ERYA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Actinobacteriae; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Pseudonocardineae; Pseudonocardiaceae;
OC Saccharopolyspora.
OX NCBI_TaxID=1836;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91220065; PubMed=2024119;
RA Donadio S., Steaver M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT "Modular organization of genes required for complex polyketide
RT biosynthesis."
RL Science 252:675-679 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RX MEDLINE=92155230; PubMed=1740151;
RA Beville D.J., Cortes J., Haydock S.F., Leadley P.F.;
RT "6-Deoxyerythronolide-B synthase 2 from saccharopolyspora erythraea.
RT Cloning of the structural gene, sequence analysis and inferred domain
RT structure of the multifunctional enzyme."
RL Eur. J. Biochem. 204:39-49 (1992).
CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
CC deoxyerythronolide B.
CC -1- COFACTOR: NADP; CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETHEINES.
CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
CC BIOSYNTHESIS.
CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH
CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3
CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
CC SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS
CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
CC RESPECTIVELY.
CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),
CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER
CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
CC OF THE FULL-LENGTH CHAIN.
CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63677; AAA26494.1; -
DR EMBL; X62569; CAA44448.1; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002085; Adh zn family.
DR InterPro; IPR000794; Ketocacyl-synt.
DR InterPro; IPR003880; Pantane_attach.
DR Pfam; PF00107; adh_zinc; 1.
DR Pfam; PF00109; ketocacyl-synt; 2.
DR Pfam; PF00550; pp-binding; 2.
DR Pfam; PF00698; Acyl transf; 2.
DR Pfam; PF02801; ketocacyl-synt C; 2.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS50075; ACP DOMAIN; 2.
KM Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
KM Phosphopantetheine; Multifunctional enzyme.
FT DOMAIN 1 1484
FT DOMAIN 1485 3567
FT DOMAIN 27 488
FT DOMAIN 559 884
FT DOMAIN 1130 1301
FT DOMAIN 1397 1467
FT DOMAIN 1485 1943
FT DOMAIN 2013 2336
FT DOMAIN 2383 3066
FT DOMAIN 3139 3322
FT DOMAIN 3415 3485
FT ACT_SITE 202 202
FT ACT_SITE 651 651
FT BINDING 1430 1430
FT ACT_SITE 1661 1661
FT ACT_SITE 2115 2115
FT NP_BIND 2961 2978
FT NP_BIND 3142 3157
FT BINDING 3448 3448
FT CONFLICT 438 438
FT CONFLICT 480 480
FT CONFLICT 1241 1241
FT CONFLICT 2664 2664
SQ SEQUENCE 3567 AA; 374413 MW; EE6284F4738AA0C0 CRC64;
Query Match 32.3%; Score 246.5; DB 1; Length 3567;
Best Local Similarity 39.0%; Pred. No. 5e-13;
Matches 57; Conservative 24; Mismatches 58; Indels 7; Gaps 3;
QY 1 YMLESTPGAGNVGSAAGDPTHEPLGATLELATDGGALLAGRLSLRHPMLADHAYGVT 60
DB 2357 FWPFPNRP--ADVSAALGVGRGHEPLLAADVPGHGVFTGRSTBQPLAEHVVGGR 2414
QY 61 VLSGATFLELATAGTYGCDRVDELTLHAPLVVVDGVSVOGVAAADGEGRLVSV 120
DB 2415 TLVGVSTLVDPALAAAGDVGLPVLELVLRPVLV-AGAGALLMSVGAPEBSGRRTIDV 2473
QY 121 YARGSGACGGGAGSGVWTCHASGYL 146
DB 2474 HAAEDVA-----DLADAQMSQHATGTL 2495
RESULT 2
MCAS_MYCBO STANDARD; PRT; 2110 AA.
AC Q02251;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Myceroctic acid synthase.
GN MAS.
OS Mycobacterium bovis.

CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CC NCBI_TaxID=1765;
CC -----
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=9240687; PubMed=1527058;
RA Mathur M., Kolatukudy P.E.;
RT "Molecular cloning and sequencing of the gene for myceroctic acid
RT synthase, a novel fatty acid elongating multifunctional enzyme, from
RT Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guerin.",
RL J. Biol. Chem. 267:19388-19395 (1992).
CC -1- FUNCTION: CATALYZES THE ELONGATION OF N-FATTY ACRYL-COA WITH
CC METHYLALANONYL-COA (NOT MALONYL-COA) AS THE ELONGATING AGENT TO
CC FORM MYCOCEROSYL LIPIDS.
CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTETHEINE.
CC -1- SUBUNIT: HOMODIMER WHOSE MONOMERS PROBABLY HAVE A HEAD TO TAIL
CC ARRANGEMENT.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -1- SIMILARITY: CONTAINS 1 ACRYL CARRIER DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M95808; AAA25369.1; -
DR PIR; B44110; B44110.
DR HSHP; P73283; 1E5M.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002085; Adh zn family.
DR InterPro; IPR000794; ketocacyl-synt.
DR InterPro; IPR003880; Pantane_attach.
DR Pfam; PF00107; adh_zinc; 1.
DR Pfam; PF00109; ketocacyl-synt; 1.
DR Pfam; PF00550; pp-binding; 1.
DR Pfam; PF00698; Acyl transf; 1.
DR Pfam; PF02801; ketocacyl-synt C; 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR PROSITE; PS50075; ACP DOMAIN; 1.
DR Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
KM Transferase; Hydrolyase; Oxidoreductase; Ligase; NADP; Membrane.
FT DOMAIN 1 430
FT DOMAIN 533 852
FT DOMAIN ? ?
FT DOMAIN ? ?
FT DOMAIN ? ?
FT ACT_SITE 2026 2096
FT ACT_SITE 177 177
FT ACT_SITE 623 623
FT NP_BIND 1561 1578
FT NP_BIND 1765 1780
FT BINDING 2059 2059
SQ SEQUENCE 2110 AA; 225577 MW; CC658215D7155300 CRC64;
Query Match 12.6%; Score 96.5; DB 1; Length 2110;
Best Local Similarity 24.3%; Pred. No. 1.7;
Matches 43; Conservative 17; Mismatches 52; Indels 65; Gaps 8;
QY 21 TEHPILGATLELATD-----GGALLAGRLSLRHPMLADHAYGVTLLSGATFLE 70
DB 894 TVHPILGSHVRLTEBEPBRHWQGDVSTVLS-----WLSDHQVHNVALLPGALYCE 944
QY 71 LALHAGTYV--GCDRVDELTLHAPLVV--PVDGVSVOV-GV-----AAADGEGRL 117
DB 945 MALAAABVGEAAVRDITPEQMLLDGEFTIDAASIDAPVNVFTVETNRNDETRH 1004
QY 118 VSYVARGG-----ACGG-----GGASGVWTCH 141
DB 1005 ATAAALRAEDDCPPPGYDITALLQAHPHAVNTRMBESFAERGVTLLGAAGLTTAH 1061

FHAB BORPE STANDARD; PRT; 3591 AA.
 ID FHAB BORPE
 AC p1225;
 RN 01-OCT-1989 (Rel. 12, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Filamentous hemagglutinin.
 GN FHAB.
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 CC Bordetella.
 CX NCBI_TaxID=520;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=90355839; PubMed=238859;
 RA Reiman D.A., Domenighini M., Tuomanen E., Rappuoli R., Falkow S.;
 RT "Genetic characterization of Bordetella pertussis filamentous
 hemagglutinin: a protein processed from an unusually large
 precursor.";
 RL Mol. Microbiol. 4:787-800(1990).
 RN [2]
 RP SEQUENCE OF 1-3261 FROM N.A.
 RX MEDLINE=89202384; PubMed=2339596;
 RA Reiman D.A., Domenighini M., Tuomanen E., Rappuoli R., Falkow S.;
 RT "Filamentous hemagglutinin of Bordetella pertussis: nucleotide
 sequence and crucial role in adherence.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:2637-2641(1989).
 CC -1- FUNCTION: EVIDENCE FOR A ROLE IN HOST-CELL BINDING AND
 CC -1- INFECTION.
 CC -1- SUBCELLULAR LOCATION: SURFACE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M60351; AAA22874.1; -;
 DR EMBL; M60351; AAA22875.1; ALT_INIT.
 DR EMBL; M60351; AAA22876.1; ALT_INIT.
 KW Antigen; Hemagglutinin.
 SQ SEQUENCE 3591 AA; 367420 MW; EF7418B3D06E5138 CRC64;
 Query Match 11.1%; Score 85; DB 1; Length 3591;
 Best Local Similarity 30.7%; Pred. No. 28;
 Matches 47; Conservative 24; Mismatches 54; Indels 28; Gaps 9;
 QY 11 GNVSA-AGLDPTETPLGATLELATDGGALLAGRLSR--SHPWADHAVG-----GT 60
 DB 792 GKVSAKSGIGLEGGAIVGAD-SLGSDDAISVSGDARVQARLADISLGAEGATLGA 850
 QY 61 VLSGATFLELALHAGTYVGCVRDELTLHAPLVVPVVGVSVOGVAAADG-----EGR 115
 DB 851 VEAAGS-----IDVAGSGSTVAAN-----SLHNRDVRVSGKAVATVATSGGLHVSGR 901
 QY 116 R--LVSYYARGSGACGGGAGSGVWTCASGVL 146
 DB 902 QLDIGAIVARGGALL-L-DGGAGVALQSAKASGTL 933
 RESULT 6
 A2AB PROHA STANDARD; PRT; 389 AA.
 ID A2AB PROHA
 AC O19054;
 RN 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-2B adrenergic receptor (Alpha-2B adrenoceptor) (Fragment).
 GN ADBA2B.
 OS *Procyon lotor* capensis habessinica (Abyssinian hyrax).
 OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Hyracoidea; Procaviidae; Procavia.
 CX NCBI_TaxID=9814;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97357151; PubMed=9214502;
 RA Springer M.S., Clevon G.C., Madsen O.J., de Jong W.W., Waddell V.G.,
 RA Amrine H.M., Stanhope M.J.;
 RT "Endemic African mammals shake the phylogenetic tree.";
 RL Nature 388:61-64(1997).
 CC -1- FUNCTION: ALPHA-2 ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
 CC INDUCED INHIBITION OF ADENYLYLATE CYCLASE THROUGH THE ACTION OF G
 CC PROTEINS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y12523; CAAT73123.1; -;
 DR HSSP; P29274; IMM.
 DR InterPro; IPR000276; GPCR_Rhodopn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECPT_P1_1.
 DR PROSITE; PS50262; G_PROTEIN_RECPT_P2_1.
 KW G-protein coupled receptor; Transmembrane; Multigene family;
 KM Phosphorylation; Lipoprotein; Palmitate.
 FT NON TER 1 1
 FT TRANSSEM 1 25
 FT DOMAIN 26 35
 FT TRANSSEM 37 62
 FT DOMAIN 63 72
 FT TRANSSEM 73 95
 FT DOMAIN 96 117
 FT TRANSSEM 118 140
 FT DOMAIN 141 156
 FT TRANSSEM 157 180
 FT DOMAIN 181 353
 FT TRANSSEM 354 377
 FT DOMAIN 378 386
 FT TRANSSEM 387 389
 FT DOMAIN 387 389
 FT DISULFID 72 151
 FT SITE 79 79
 FT SITE 163 163
 FT SITE 167 167
 FT NON TER 389 389
 SQ SEQUENCE 389 AA; 42528 MW; 4F2089EA140876E1 CRC64;
 Query Match 11.1%; Score 84.5; DB 1; Length 389;
 Best Local Similarity 26.3%; Pred. No. 3.2; Indels 59; Gaps 5;
 Matches 44; Conservative 8; Mismatches 56; Indels 59; Gaps 5;
 QY 8 PGAGNVSAAGLDPTETPLGATLELATDGGALLAGRLSR--SHPWADHAVG-----DHA 62
 DB 194 PRAKGAPEBESQPPH-----LHAGPLALANPTLATSLAVDEANGHSK 239
 QY 63 LSGATFLELALHAGTYVGCVRDELTLHAPLVVPVVGVSVOGVAAADG-----DGG 114
 DB 240 LTGEKERETSSEDPT-----PTLORSWPTLPSSGSGQKKGVCAGASPEEAEGBE 290
 QY 115 RRVSVYARGSGACG-----GGGAGSGVW 138
 DB 291 SRPLSVPASPASGCPHLOPQGSQVLAATLRGOVLLRGVGAAGGQW 337

DT 15-JUN-2002 (Ref. 41, last annotation update)
DE Octamer-binding transcription factor 6 (Oct-6) (POU-domain
DE transcrip-tion factor SCIP) (TST-1).
GN POU3F1 OR OTF6 OR OTF-6 OR OCT-6 OR SCIP.
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90378306; PubMed=1975954;
RA Monnik E.S., Kuhn R., Weinmaster G., Trapp B.D., Lemke G.;
RT "Expression and activity of the POU transcription factor SCIP";
RL Science 249:1300-1303(1990).
RN [2]
RP SEQUENCE OF 265-369 FROM N.A.
RX MEDLINE=69295573; PubMed=2739723;
RA He X., Treacy M.N., Simmons D.M., Ingraham H.A., Swanson L.W.,
RA Rosenthal M.G.;
RL Nature 340:662-662(1989).
RN [4]
RP SEQUENCE OF 35-451 FROM N.A.
RX MEDLINE=91141528; PubMed=1705013;
RA He X., Gerrero M.R., Simmons D.M., Park R.E., Lin C.R.,
RA Swanson L.W., Rosenthal M.G.;
RT "Tst-1, a member of the POU domain gene family, binds the promoter of
RT the gene encoding the cell surface adhesion molecule P0.";
RL Mol. Cell. Biol. 11:1739-1744(1991).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE OCTAMER MOTIF
CC ('ATTGCAT'). THOUGHT TO BE INVOLVED IN EARLY EMBRYOGENESIS AND
CC NEUROGENESIS. BOTH IN NORMAL DEVELOPMENT AND IN RESPONSE TO NERVE
CC TRANSECTION. SCIP EXPRESSION WAS TRANSIENTLY ACTIVATED ONLY
CC DURING THE PERIOD OF RAPID CELL DIVISION THAT SEPARATES THE
CC PREMYELINATING AND MYELINATING PHASES OF SCHWANN CELL
CC DIFFERENTIATION. IN COTRANSCRIPTION ASSAYS, SCIP ACTED AS A
CC TRANSCRIPTIONAL REPRESSOR OF MYELIN SPECIFIC GENES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: NEURAL TISSUES AND TESTIS.
CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
CC CLASS-3 SUBFAMILY
CC -1- SIMILARITY: CONTAINS 1 HOMEOBOX DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M72711; AAA42118.1; -;
DR EMBL: M63712; AAA42303.1; -;
DR PIR: S05044; S05044.
DR PIR: S05449; S05449.
DR HSSP: P14859; 1OCT.
DR TRANSFAC: T00969; -;
DR InterPro: IPR001356; Homeobox.
DR InterPro: IPR000327; POU_domain.
DR Pfam: PF00046; homeobox; 1.
DR Pfam: PF00157; pou; 1.
DR PRINTS: PR00028; POUDOMAIN.
DR ProDom: PD000010; Homeobox; 1.
DR ProDom: PD000583; POU_domain; 1.
DR SMART: SM00389; HOX; 1.
DR SMART: SM00352; POU; 1.
DR PROSITE: PS00027; HOMEOBOX_1; 1.

DR PROSITE; PS00035; POU 1; 1.
 DR PROSITE; PS00465; POU 2; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation; Repressor;
 KM Homeobox.
 FT DOMAIN 251 321 POU.
 FT DNA_BIND 339 398 HOMEBOX.
 FT CONFLICT 423 423 S->G (IN REF. 4).
 SQ SEQUENCE 451 AA; 45496 MW; E32626AV59125EAA CRC64;
 Query Match 11.0%; Score 84; DB 1; Length 451;
 Best Local Similarity 26.7%; Pred. No. 4.1;
 Matches 47; Conservative 9; Mismatches 76; Indels 44; Gaps 8;
 QY 8 PGAGNVSAAGLDPTHEPLGATLELATDGGALLAGRL-----SLRSHPMW---ADHAV-- 57
 DB 12 PGGG---AGTGCPMLHPDADAAAAAERLHGAAYREVOKLMHEHWLGAGCHGVGL 68
 QY 58 -----GGTVLLSGATFLLELALHAGTVVGCDDRYDE---LTLHAPLV----- 94
 DB 69 AHPQWLPTGGGGGDMWAGPHLEHGKAGGGSTG--RADDDGGGGGFHARLVHGAHAGA 126
 QY 95 VPDGGSVVOYGVV-----AADGGRRLVSIVYARGSGACGGGASGGVWTCASG 144
 DB 127 AWAQGGTAHHLGPAMSPSPGAGGCHOPPLGLVQAAYVPGGGGGGLAGMLAAGGGG 182
 RESULT 9
 OCT6_MOUSE STANDARD; PRT; 449 AA.
 AC P21952;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Octamer-binding transcription factor 6 (OCT-6) (POU-domain
 DE transcription factor SCIP).
 GN POU3F1 OR OTF-6 OR OCT6 OR SCIP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Brain;
 RX MEDLINE=91081324; PubMed=1979677;
 RA Meijer D., Graus A., Kraay R., Langveld A., Mulder M.P., Grosveld G.;
 RT "The octamer binding factor Oct6: cDNA cloning and expression in
 RT early embryonic cells.";
 RL Nucleic Acids Res. 18:7357-7365 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=91006074; PubMed=1976514;
 RA Suzuki N., Rohdewold H., Neuman T., Gruss P., Schoeler H.R.;
 RT "Oct-6: a POU transcription factor expressed in embryonal stem cells
 RT and in the developing brain.";
 RL EMBO J. 9:3723-3732 (1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91204456; PubMed=1840678;
 RA Zimmerman E.C., Jones C.M., Fet V., Hogan B.L.M., Magnuson M.A.;
 RT "Nucleotide sequence of mouse SCIP cDNA, a POU-domain transcription
 RT factor.";
 RL Nucleic Acids Res. 19:956-956 (1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92228768; PubMed=1565620;
 RA Hara Y., Rovescelli C., Kim Y., Nirenberg M.;
 RT "Structure and evolution of four POU domain genes expressed in mouse
 RT brain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:3280-3284 (1992).
 CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE OCTAMER MOTIF
 CC ('ATTGCAAT'). THOUGHT TO BE INVOLVED IN EARLY EMBRYOGENESIS AND

CC NEUROGENESIS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONAL STEM CELLS AND IN THE
 CC DEVELOPING BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
 CC -1- CLASS-3 SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X54628; CAA38445.1; -;
 DR EMBL; X56959; CAA40280.1; -;
 DR EMBL; X57482; CAA40720.1; -;
 DR EMBL; M88302; AAA39963.1; -;
 DR PIR; S11999; S11999.
 DR PIR; S13083; S13083.
 DR PIR; S30205; S30205.
 DR PIR; S31226; S31226.
 DR HSSP; P14859; 1OCT.
 DR TRANSPAC; T00656; -;
 DR MGD; MGI:101896; Pou3f1.
 DR InterPro; IPR001355; Homeobox.
 DR InterPro; IPR000327; POU_domain.
 DR Pfam; PF00046; homeobox; 1.
 DR Pfam; PF00157; pou; 1.
 DR PRINTS; PR00028; POUDOMAIN.
 DR Prodom; PD000010; Homeobox; 1.
 DR Prodom; PD000583; POU domain; 1.
 DR SMART; SM00389; HOX; 1.
 DR SMART; SM00352; POU; 1.
 DR PROSITE; PS00027; HOMEBOX 1; 1.
 DR PROSITE; PS00035; POU 1; 1.
 DR PROSITE; PS00465; POU 2; 1.
 DR PROSITE; PS00071; HOMEBOX 2; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation; Homeobox.
 KM DOMAIN 249 319 POU.
 FT DNA_BIND 337 396 HOMEBOX.
 FT CONFLICT 35 35 MISSING (IN REF. 2).
 SQ SEQUENCE 449 AA; 45323 MW; 82DD1335A0C4F43C CRC64;
 Query Match 10.3%; Score 83; DB 1; Length 449;
 Best Local Similarity 26.4%; Pred. No. 4.9;
 Matches 46; Conservative 10; Mismatches 76; Indels 42; Gaps 8;
 QY 8 PGAGNVSAAGLDPTHEPLGATLELATD---GGALLAGRLSLRSHPMW---ADHAV----- 57
 DB 12 PGGG---AGTGCPMLHPDADAAAAAERLHGAAYREVOKLMHEHWLGAGCHGVGLAH 68
 QY 58 -----GGTVLLSGATFLLELALHAGTVVGCDDRYDE---LTLHAPLV-----VP 96
 DB 69 PQWLPTGGGGGDMWAGPHLEHGKAGGG--GTRADDDGGGGGFHARLVHGAHAGAAM 126
 QY 97 VPDGGSVVOYGVV-----AADGGRRLVSIVYARGSGACGGGASGGVWTCASG 144
 DB 127 AQAQGGTAHHLGPAMSPSPGAGGCHOPPLGLVQAAYVPGGGGGGLAGMLAAGGGG 180
 RESULT 10
 FABH_RHOCA STANDARD; PRT; 324 AA.
 ID FABH_RHOCA
 AC P30790;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 3-oxoacyl-l-acyl-carrier-protein synthase III (EC 2.3.1.41) (Beta-
 DE ketoacyl-ACP synthase III) (KAS III).
 GN FABH.

QY 71 -LALHAGTYVCGDVRDELTLHAPLVV-----PVDG-GVSVGVVAAADGEGRRLL 117
 DB 79 ALTAGAGSYAAEAASAPLEGVLDVINAAPALLGRLLIONGANAPGTANGDGGILL 138
 QY 118 VSVYARGGSAC-----GGGAGAG 135
 DB 139 IGGAGAGSGAAGMPGGGGAAG 161

RESULT 12

WA22 MYCTU STANDARD; PRT; 914 AA.
 AC 006794;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-OCT-2001 (Rel. 40, Last annotation update)
 DE WAG22 antigen precursor.
 GN WAG22 OR RV1759C OR MT1807 OR MTCY28.25C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinobacteriales; Corynebacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=H37RV;
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
 RL Nature 393:537-544 (1998).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=CDC 1551 / Oshkosh;
 RC Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PGSS SUBFAMILY.
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A FRAMESHIFT IN POSITION 85.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z95890; CAB09322.1; -;
 DR EMBL; AB007040; -; NOT_ANNOTATED_CDS.
 DR TIGR; MT1807; -;
 DR TubercuList; RV1759C; -;
 DR InterPro; IPR000084; PE_region.
 DR Pfam; PF00934; PE; 1.
 KW Antigen; Repeat; Signal; Complete proteome.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 914 WAG22 ANTIGEN.
 SQ SEQUENCE 914 AA; 74354 MW; F6953C3DBE8B6AC8 CRC64;

Query Match 10.5%; Score 80.5; DB 1; Length 914;

Best Local Similarity 27.9%; Pred. No. 17;
 Matches 38; Conservative 9; Mismatches 60; Indels 29; Gaps 4;

QY 9 GAGVNSAGADDPHEPLLTGTLTDLADGALLGRSLRSHPEPLADHAGVGTLLSGATF 68
 DB 557 GAGGAGAGAGTNGAGAGAGAGGILFTGTGAGAGAGVTA-----GAGGAGAGAGSAP 608

QY 69 LELAHAGTYVCGDVRDELTLHAPLVVPVGVGVVAAADGEGRRLVSVYARGGSAC 128
 DB 609 L-----IGSGGTG-----VGGAATTGGVGGAGAGNGLLGAAGLGG--C 647

QY 129 GGGGAGSGVWTHASG 144
 DB 648 GCGAFTAGVTTGAGAG 663

RESULT 13

MB48 MYCTU STANDARD; PRT; 460 AA.
 AC 0933R8;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Antigen MTB48.
 GN MTB48 OR RV3881C OR MT3996 OR MTW027.16C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinobacteriales; Corynebacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Erdmann;
 RC MEDLINE=2130746; PubMed=11427558;
 RA Iodes M.J., Dillon D.C., Mohamath R., Day C.H., Benson D.R., Reynolds L.D., McKell P., Sampaio D.P., Skelky Y.A.W., Badaro R., Persing D.H., Reed S.G., Houghton R.L.;
 RA "Serological expression cloning and immunological evaluation of MTB48, a novel Mycobacterium tuberculosis antigen."
 RL J. Clin. Microbiol. 39:2485-2493 (2001).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=H37RV;
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
 RL Nature 393:537-544 (1998).
 [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=CDC 1551 / Oshkosh;
 RC Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. A processed form is shed into the medium during culture.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial


```

DR EMBL: AL583925; CAC31968.1;
DR Leproma; ML2451;
DR InterPro; IPR002915; Deoc.
DR InterPro; IPR003009; FMN_enzyme.
DR Pfam; PF01791; Deoc, 1.
KW Lyase; Schiff base; Complete proteome.
FT BINDING 161 SCHIFF-BASE (BY SIMILARITY).
SQ SEQUENCE 226 AA; 22527 MM; 2B79A2D09B84F8A4 CRC64;

Query Match 10.1%; Score 77.5; DB 1; Length 226;
Best Local Similarity 27.0%; Pred. No. 7.2;
Matches 44; Conservative 15; Mismatches 59; Indels 45; Gaps 8;

Qy 5 STQDGAQNVAA---GLDPTIEHPL-----GAT-LELATGCGALLAGRL-SLR 47
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 53 AVEAGAGAGLVAATVAVGPPSGEHSIVYKQAEVAQVAGSGATEIDIVIAVGAALADDLAVR 112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 48 SHPLADHAYAGTVL---LSGATFLELA-----LHAGYVGCDAVDELTLHAPLVVP 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 113 SDIEWVRSALAGVVLKTVISAVALLGGANERTLAIEGVHVAEACADPVYKSTGHFPA--- 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 97 VDGVSIVQVGVAAADGEGRLVSVYARGSGACGGGSGGVT 139
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 170 --GGASVRAVTLMTMEAVGRL-----GVKASGGIRT 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
ID ym05_MYCTU STANDARD; PRT; 358 AA.
AC O10354;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein Rv2205C.
GN Rv2205C OR MT2261 OR MTCY190.16C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteriae (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=96295987; Pubmed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagals K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R., Squares R.,
RA Rutsen J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deloy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolova M.D., Salzberg S.L.,
RA DeJongh A., Ucterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:55:32 ; Search time 3.26017 Seconds
(without alignments)
4305.183 Million cell updates/sec

Title: US-09-914-286-3_COPY_2983_3128

Perfect score: 764
Sequence: 1 YWLESTQPGAGNVSAAGLDPE.....ACGGGAGSGGWTCHASGVL 146

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	385.5	50.5	3739	2 T17410	polyketide synthas
2	361	47.3	6420	2 T30283	polyketide synthas
3	346.5	45.4	2103	2 G86925	probable polyketid
4	332	43.5	5069	2 T17464	rifamycin polyketi
5	329	43.1	10223	2 T30225	polyketide synthas
6	324.5	42.5	4151	2 G70944	probable polyketid
7	320	41.9	4613	2 T17409	polyketide synthas
8	317	41.5	7576	2 T17428	PKS06 polyketide s
9	315.5	41.3	1937	2 T03224	Probable polyketid
10	314.5	41.2	8563	2 T30226	polyketide synthas
11	313.5	41.0	2124	2 T28658	polyketide synthas
12	312	40.8	2126	2 H70621	probable polyketid
13	308.5	40.4	3413	2 T17467	rifamycin polyketi
14	304	39.8	1616	2 G70668	polyketide synthas
15	300	39.3	1763	2 T17465	rifamycin polyketi
16	299	39.1	1602	2 H70984	probable polyketid
17	279.5	36.6	6260	2 T30228	polyketide synthas
18	271.5	35.5	1762	2 T03222	probable polyketid
19	245.5	32.1	2100	2 T03223	probable polyketid
20	240.5	31.5	2723	2 T03221	probable polyketid
21	234	30.6	3573	2 S23070	erythronolide synt
22	201.5	26.4	1728	2 T17466	rifamycin polyketi
23	147.5	19.3	4735	2 A70984	rifamycin polyketi
24	126.5	16.6	2188	2 B70984	probable polyketid
25	121	15.8	1827	2 B70984	probable polyketid
26	103	13.5	2201	2 S73014	polyketide synthas
27	101.5	13.3	2108	2 H70819	probable polyketid
28	99.5	13.0	2118	2 S72705	mycocerosate synth
29	97.5	12.8	2111	2 A70668	mycocerosate synth

30	97.5	12.8	2126	2 E70522	probable polyketid
31	96.5	12.6	2110	2 B44110	mycocerosate synth
32	95	12.4	741	2 G70917	hypothetical glyci
33	95	12.4	2116	2 C86926	probable mycoceros
34	92	12.0	1582	2 E70876	probable polyketid
35	91.5	12.0	2518	2 A12140	polyketide synthas
36	91	11.9	1822	2 S73021	polyketide synthas
37	91	11.9	1822	2 F87203	polyketide synthas
38	89	11.6	433	2 F75566	probable benzozate
39	88.5	11.6	775	2 T07172	subtilisin-like pr
40	88	11.5	2174	2 E95965	hypothetical glyci
41	88	11.5	2554	2 AB3528	extremophilic seri
42	87.5	11.5	1028	2 T03516	probable outer mem
43	86.5	11.3	509	2 T34871	probable membrane
44	86	11.3	1027	2 F87370	alpha-L-rhamnoside
45	86	11.3	4936	2 AH2515	hypothetical prote

ALIGNMENTS

RESULT 1
T17410
polyketide synthase type I - Streptomyces venezuelae
C:Species: Streptomyces venezuelae
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
C:Accession: T17410
R:Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
A:Title: A gene cluster for macroide antibiotic biosynthesis in streptomyces venezuelae
A:Reference number: Z18773; PMID:98445333; PMID:9770448
A:Accession: T17410
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 13739 <XUE>
A:Cross-references: EMBL:AF079138; NID:G3808326; PID:G3800835; PIDN:AA669330.1
C:Gene(s):
A:Gene: p1KAI1
C:Superfamily: acyl carrier protein homology
C:Keywords: antibiotic biosynthesis; carrier protein
F:1445-1516/Domain: acyl carrier protein homology <ACP1>
F:3570-3641/Domain: acyl carrier protein homology <ACP2>

Query Match 50.5%; Score 385.5; DB 2; Length 3739;
Best Local Similarity 54.8%; Pred. No. 5,3e-25;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YWLESTQPGAGNVSAAGLDPEHPLGATLELATDGAALAGRLSRSHEMLADHAGGT 60
DB 2406 YWQPDLSAAGDITSGAAGHPLGAVALADSDGCLTGSLSRTHEMLADHAGGT 2465
QY 61 VLSGATLELAHAGTYCCDVRDELTLHAPLVFVVDGVSQVGAADSGRRLVSV 120
DB 2466 VLLPGTAFVFLARADQVCCDLVEBLTLDAPLVLRGAVRVOLSGVSDSGRTFGL 2525
QY 121 YARGSGAGGAGSGGWTCHASGVL 146
DB 2526 YAHPEADPG-----EAEWTRHATGVL 2546

RESULT 2
T30283
polyketide synthase - Streptomyces sp. (strain MA6548)
C:Species: Streptomyces sp.
A:Variety: strain MA6548
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
R:McNameid, H.; Cai, S.J.; Shafiee, S.J.; Billston, K.O.
Eur. J. Biochem. 244, 74-80, 1997
A:Title: Structural organization of a multifunctional polyketide synthase involved in c
A:Reference number: Z20806; PMID:97217427; PMID:9063448
A:Accession: T30283
A:Status: preliminary; translated from GB/EMBL/DBJ

	Query Match	41.5%	Score 317;	DB 2;	Length 7576;
	Best Local Similarity	47.6%;	Pred. 9.8e-19;		
	Matches	70;	Conservative 14;	Mismatches 49;	Indels 14; Gaps 4;
Qy	1 YWLESTPGAG-NVSAAGIDPTENHPLGATTELATDGGALLAGRLSLRSHFWLADHAVGG	59			
Db	6851 YWIR---PATGDASGLGDANDHPDLAASVALPDSDEVLLTGRSLATHFHWLGHTYDG	6907			
Qy	60 TVLLSGATFLALAHAGTYVGCDRVDELTLHAFLVPVPDGVSVQVGAADDGSRRLVS	119			
Db	6908 EVLLPFGPVFEVLACGADBEARDLDLDELVIETPLPALPATGAVQVRVTAAADDAGRRAYR	6967			
Qy	120 VYARGGSACGGGGGASGG-VMTCHASGV	145			
Db	6968 IHAR-----AHGGRWTIRHASGT	6985			

[illegible]

C:Superfamily 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:54-503/Domain: acetate-CoA ligase homology <AC>
F:1339-1724/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F:1817-2091/Domain: [acyl-carrier-protein]-S-malonyltransferase homology <AMT1>
F:3159-3270/Domain: acyl carrier protein homology <ACP>
F:3314-3706/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F:4787-4858/Domain: acyl carrier protein homology <ACP>
F:4902-5293/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F:5386-5659/Domain: [acyl-carrier-protein]-S-malonyltransferase homology <AMT2>
F:6760-6831/Domain: acyl carrier protein homology <ACP>
F:6875-7269/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F:7362-7636/Domain: [acyl-carrier-protein]-S-malonyltransferase homology <AMT3>
F:8432-8483/Domain: acyl carrier protein homology <ACP>
F:84822/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match	41.2%	Score 314.5	DB 2	Length 8563
Best Local Similarity	44.5%	Pred. No. 1.8e-18		
Matches	65	Conservative 18	Mismatches 46	Indels 17
			Gaps	2

QY	1	YMLESTPGAGNVSAGLDPTHEPLGATTELTGAGLLAGRLSRHPLADHAVG	60
DB	7692	YMWVSVDRAAD-----GHLGAVVELPESDCVLLTGRVSLATHAMLDHAWMR	7742
QY	61	VLLSGATFLLELTHAGTYVGCGRYDELTLTHPLVVPVPGGSVVOYGAAADGSRRLVS	120
DB	7743	VLLPGLTFLVELVYHMAAGEGCDVDVDELVIETPLLLPOTGVOLSVSGEADSGHRVVT	7802
QY	121	YARGSAACGGGAGCGVWTCHASGL	146
DB	7803	FSRADN-----DTWIRHVSATV	7820

```

RESULT 11
T28658
polyketide synthase - Sorangium cellulosum (fragment)
C|Species: Sorangium cellulosum
C|Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 01-Dec-2000
C|Accession: T28658
R|Schupp, T.; Toupet, C.; Cluzel, B.; Neff, S.; Hill, S.; Beck, J.J.; Ligon, J.M.
J|Bacteriol. 177, 3673-3679, 1995
A|Title: A Sorangium cellulosum (Myxobacterium) gene cluster for the biosynthesis of th
inomyces.
A|Reference number: Z20499; MUID:95325306; PMID:7601830
A|Accession: T28658
A|Status: preliminary; translated from GB/EMBL/DDBJ
A|Molecule type: DNA
A|Residues: 1-2124 <SCH>
A|Cross-references: EMBL:U24241; NID:G868150; PID:G868151; PIDN:AAA79984.1
A|Experimental source: strain So ce26
C|Function:
A|Description: involved in soraphen A biosynthesis
C|Superfamily: acyl carrier protein homology
C|Keywords: carrier protein; phosphopantetheine; phosphoprotein
F|335-406/Domain: acyl carrier protein homology <ACP>
F|370/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match          41.0%; Score 313.5; DB 2; Length 2124;
Best Local Similarity 45.2%; Pred. No. 5.1e-19;
Matches 66; Conservative 20; Mismatches 55; Indels 5; Gaps 1;

QY      1 YWLESTOPAGNVSAAGIDPTENHPLIGATLELDGDGALLGRSLRSHPWLADHAVGT 60
       :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
DB     1314 FWLDASTHADVDVASAGITSDAPHLGLAAVALADRDGVFTGRSLASHPWLDEHVVGII 1373
QY      61 VLLSGATLEALACGTVGGCDRYDELTLHA PLVPVNDGGSVYGVAAGAEGRRILSV 120
       :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
DB     1374 PCPARRLLETALHVAHLVGDIVEDVTLLDPLPALPSGAVALLOISVGPADGAGRRAI SV 1433
QY      121 YARGSACGGGASGAVTCASGVL 146
       :::: | :||| |||
DB     1434 HSRRHDA-----LDDGPTWRHASGSL 1454

```


A;Reference number: Z18802

A;Accession: T17465
A;Status: preliminary

A;Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A;Molecule type: DNA
A;Residues: 1-1763

A;Residues: 1-1763 <SCH>
A:Cross-references: EMBL.

A/Cross-references: EMBL:AJ223012; NID:e1227119; PID:e1227122; PIDN:CA11037.1
A/Experimental source: strain PBG A3136

A;Experimental source: strain LBG A3136
C;Superfamily: Streptomyces hygroscopicus

C/Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoacyl:acyl-carrier-protein S-malonyltransferase homology

C;Keywords: carrier protein
F;1615-1686/Domain: acyl carrier protein homology <ACP>

F;1615-1686/Domain: acyl carrier protein homology <ACP>

Query Match	39.3%;	Score 300;	DB 2;	Length 1763;
Best Local Similarity	43.5%;	Bred No. 6	10-10.	

Best Local Similarity 43.5%; Pred. No. 6.1e-18;
Matches 64; Conservative 19; Mismatches 50

Matches 64; Conservative 19; Mismatches 50; Indels 14; Gaps 3;

QY 1 YWLESTQPGA-GNVSAGLDPTTEHPLLGATLELATDGGALLAGRLSLRSHPWADHAVGG 59

Db 891 FWL---RPAAGADAVSLGQAAAEHPLGLGAVLRPQSDGLVFTSRSLRTNHPMLADHPTGG 947

Qy 60 TVLLSGATFLELTHAGTYVGC DRVDELTHAPLVVPV DGVSVQYGVAAADGEGRLVS 119

Db 948 VLFPGTGLVELAVRAGDEAGCPVLDELVTETAPLVPGGGVNVQTVTSGPDQNGLRTVD 1007

QY 120 VYARGGSA CGGGASGVWTC HASGVL 146

```

Db      1008 IHSQ-----RDDVWTRHATGV 1024
          ::::| | | | | : | :

```

Search completed: June 17, 2003, 13:12:24
Job time : 4 36017 secs

Job time : 4.26017 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 13:08:02 ; Search time 4.54907 Seconds
(without alignments)
3431.399 Million cell updates/sec

Title: US-09-914-286-3_COPY_2983_3128

Perfect score: 764

Sequence: 1 YWLESTQPGAGNVSAAGLDP.....ACGGGAGSGGWTCHASGVL 146

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB_pdp:*
2: /cgn2_6/ptodata/1/pubppaa/PTCT_NEW_PUB_pdp:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB_pdp:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB_pdp:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB_pdp:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB_pdp:*
7: /cgn2_6/ptodata/1/pubppaa/PTCUS_PUBCOMB_pdp:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB_pdp:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB_pdp:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB_pdp:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB_pdp:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB_pdp:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB_pdp:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB_pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	385.5	50.5	3739	9 US-09-860-846-33	Sequence 33, App1
2	385.5	50.5	3739	9 US-09-988-384B-33	Sequence 33, App1
3	385.5	50.5	3739	9 US-09-836-821-33	Sequence 33, App1
4	385.5	50.5	3739	9 US-09-793-708-2	Sequence 2, App1
5	385.5	50.5	3739	10 US-09-861-289-33	Sequence 33, App1
6	385.5	50.5	11877	9 US-09-860-846-6	Sequence 6, App1
7	385.5	50.5	11877	9 US-09-836-821-6	Sequence 6, App1
8	385.5	50.5	11877	10 US-09-861-289-6	Sequence 6, App1
9	385.5	50.5	12199	9 US-09-988-384B-6	Sequence 6, App1
10	324	42.4	5215	9 US-09-860-846-2	Sequence 2, App1
11	324	42.4	5215	9 US-09-988-384B-2	Sequence 2, App1
12	324	42.4	5215	9 US-09-836-821-2	Sequence 2, App1
13	324	42.4	5215	10 US-09-861-289-2	Sequence 2, App1
14	320	41.9	4551	9 US-09-793-708-1	Sequence 1, App1
15	320	41.9	4613	9 US-09-860-846-31	Sequence 31, App1
16	320	41.9	4613	9 US-09-988-384B-31	Sequence 31, App1
17	320	41.9	4613	9 US-09-836-821-31	Sequence 31, App1
18	320	41.9	4613	10 US-09-861-289-31	Sequence 31, App1
19	319	41.8	3816	9 US-09-808-880-3	Sequence 3, App1

20	304	39.8	1616	9 US-09-712-363-262	Sequence 262, App
21	155	20.3	1832	9 US-10-014-717-4	Sequence 4, App1
22	155	20.3	7257	9 US-10-014-717-5	Sequence 5, App1
23	154	20.2	3798	9 US-10-014-717-6	Sequence 6, App1
24	147	19.2	2439	9 US-10-014-717-7	Sequence 7, App1
25	121	15.8	1827	9 US-09-712-363-261	Sequence 261, App
26	118.5	15.5	2563	9 US-09-836-705-46	Sequence 46, App1
27	96	12.6	1674	9 US-10-123-155-453	Sequence 453, App
28	85	11.1	256	9 US-09-880-748-1607	Sequence 1607, App
29	84.5	11.1	423	10 US-09-073-009-142	Sequence 142, App
30	84.5	11.1	423	10 US-09-793-306-142	Sequence 142, App
31	84.5	11.1	462	9 US-10-270-333-33	Sequence 33, App1
32	84.5	11.1	10	9 US-09-287-849-16	Sequence 16, App1
33	84.5	11.1	856	10 US-09-287-849-112	Sequence 12, App1
34	83	10.9	1567	9 US-10-123-155-215	Sequence 215, App
35	82.5	10.8	1797	9 US-10-184-634-331	Sequence 331, App
36	82.5	10.8	1797	9 US-10-184-634-331	Sequence 331, App
37	81.5	10.7	1094	9 US-10-184-644-387	Sequence 387, App
38	81.5	10.7	1094	9 US-10-184-634-387	Sequence 387, App
39	81.5	10.7	1621	9 US-10-184-644-145	Sequence 145, App
40	81.5	10.7	1621	9 US-10-184-634-145	Sequence 145, App
41	80.5	10.5	2698	9 US-10-123-155-25	Sequence 25, App1
42	80.5	10.5	4374	9 US-10-123-155-125	Sequence 125, App
43	79.5	10.4	251	9 US-09-880-748-1510	Sequence 1510, App
44	79.5	10.4	1264	9 US-10-123-155-1	Sequence 1, App1
45	79.5	10.4	1857	9 US-10-123-155-365	Sequence 365, App

ALIGNMENTS

RESULT 1
US-09-860-846-33 Application US/09860846
Sequence 33, Application US/09860846
Patient No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, Y.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438051
CURRENT APPLICATION NUMBER: US/09/860,846
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-860-846-33

Query Match 50.5%; Score 385.5; DB 9; Length 3739;
Best Local Similarity 54.8%; Pred. No. 4.2e-26;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YWLESTQPGAGNVSAAGLDPTFHPPLGATLELATGCGALLAGRLSRHPMLADHAGVT 60
DB 2406 YWQPDLSAAGDITTSAGLAHEHPLGLAAVALADSGCLLTGSLSTRHPMLADHAGVT 2465
QY 61 VILSGATFLELHAGTYGCDVDELTTLAIPVPPDGVSVQVGAADDEGRRLVSV 120
DB 2466 VILPGTAFVELARADQVGCDDVBEUTLDAPLPRGAVRVQLSVGASDESGRRTFGL 2525
QY 121 YARGSGACGGGAGSGGWTCHASGVL 146
DB 2526 YAHPEDAFG-----EAEMTRHATGVL 2546

RESULT 2
US-09-988-384B-33

```
; Sequence 33, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 33
; LENGTH: 3739
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-33
```

```
Query Match          50.5%; Score 385.5; DB 9; Length 3739;
Best Local Similarity 54.8%; Pred. No. 4.2e-26;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;
```

```
QY 1 YMLSTPGAGNVSAGLDPTHEHPLTGATLELATDGGALLAGRLSLRSHPLADHVAAGT 60
DB 2406 YMPDPDISAAGDITTSAGLGAHEHPLGAAVALADSDGLTGSLSLRTHPWLADHVAAGT 2465
QY 61 VLLSGATFLELALHAGTYVGCDDRYDELTLHAPLVVPVDSGVSVGVAAADGEGRLVSV 120
DB 2466 VLLPGTAFVELAFRAGDQVCDLVEBELTLDAPIVLPFRGAVRVQLSVGASDSGRRTFGL 2525
QY 121 YARGSGACGGGASGGVWTCASGVL 146
DB 2526 YAHPEDAFG-----EAETRHATGVL 2546
```

```
RESULT 3
US-09-836-821-33
; Sequence 33, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 3739
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-836-821-33
```

```
Query Match          50.5%; Score 385.5; DB 9; Length 3739;
Best Local Similarity 54.8%; Pred. No. 4.2e-26;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;
```

```
QY 1 YMLSTPGAGNVSAGLDPTHEHPLTGATLELATDGGALLAGRLSLRSHPLADHVAAGT 60
DB 2406 YMPDPDISAAGDITTSAGLGAHEHPLGAAVALADSDGLTGSLSLRTHPWLADHVAAGT 2465
QY 61 VLLSGATFLELALHAGTYVGCDDRYDELTLHAPLVVPVDSGVSVGVAAADGEGRLVSV 120
DB 2466 VLLPGTAFVELAFRAGDQVCDLVEBELTLDAPIVLPFRGAVRVQLSVGASDSGRRTFGL 2525
```

```
QY 121 YARGSGACGGGASGGVWTCASGVL 146
DB 2526 YAHPEDAFG-----EAETRHATGVL 2546
```

```
RESULT 4
US-09-793-708-2
; Sequence 2, Application US/09793708
; Publication No. US20030104597A1
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002121
; CURRENT APPLICATION NUMBER: US/09/793,708
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 09/657,440
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: US 09/141,908
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: US 09/073,538
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: US 08/846,247
; PRIOR FILING DATE: 1997-04-30
; PRIOR APPLICATION NUMBER: US 60/134,990
; PRIOR FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 3739
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-793-708-2
```

```
Query Match          50.5%; Score 385.5; DB 9; Length 3739;
Best Local Similarity 54.8%; Pred. No. 4.2e-26;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;
```

```
QY 1 YMLSTPGAGNVSAGLDPTHEHPLTGATLELATDGGALLAGRLSLRSHPLADHVAAGT 60
DB 2406 YMPDPDISAAGDITTSAGLGAHEHPLGAAVALADSDGLTGSLSLRTHPWLADHVAAGT 2465
QY 61 VLLSGATFLELALHAGTYVGCDDRYDELTLHAPLVVPVDSGVSVGVAAADGEGRLVSV 120
DB 2466 VLLPGTAFVELAFRAGDQVCDLVEBELTLDAPIVLPFRGAVRVQLSVGASDSGRRTFGL 2525
QY 121 YARGSGACGGGASGGVWTCASGVL 146
DB 2526 YAHPEDAFG-----EAETRHATGVL 2546
```

```
RESULT 5
US-09-861-289-33
; Sequence 33, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
```

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRF
ORGANISM: Streptomyces venezuelae
US-09-861-289-33

Query Match 50.5%; Score 385.5; DB 10; Length 3739;
Best Local Similarity 54.8%; Pred. No. 4.2e-26;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YMLESTPGAGNVSAAGLDPTHEPLLGATLELATDGGALLAGRLSLRSHPMADHAGVT 60
DB 2406 YWQPDLSAAGDITTSAGLGAHEHPLGAAVALADSDCLLTGSLRTHPMADHAGVT 2465
QY 61 VLLSGATFLELALHAGTYGCDRVDELTLHAPLVVPVDGVSVOGVAADGEGRLVSV 120
DB 2466 VLLPGTAFVELAFRAADQVGCDELVEELTLDAPLVLPFRGAVRVOLSVGASDESGRRTFGL 2525
QY 121 YARGSGACGGGASGGVWTCASGVL 146
DB 2526 YAHPEADAPG-----EAEWTRHATGVL 2546

RESULT 6
US-09-860-846-6
Sequence 6, Application US/09860846
Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/860.846
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105.537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRF
ORGANISM: Streptomyces venezuelae
US-09-860-846-6

Query Match 50.5%; Score 385.5; DB 9; Length 11877;
Best Local Similarity 54.8%; Pred. No. 1.6e-25;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YMLESTPGAGNVSAAGLDPTHEPLLGATLELATDGGALLAGRLSLRSHPMADHAGVT 60
DB 7355 YWQPDLSAAGDITTSAGLGAHEHPLGAAVALADSDCLLTGSLRTHPMADHAGVT 7414
QY 61 VLLSGATFLELALHAGTYGCDRVDELTLHAPLVVPVDGVSVOGVAADGEGRLVSV 120
DB 7415 VLLPGTAFVELAFRAADQVGCDELVEELTLDAPLVLPFRGAVRVOLSVGASDESGRRTFGL 7474
QY 121 YARGSGACGGGASGGVWTCASGVL 146
DB 7475 YAHPEADAPG-----EAEWTRHATGVL 7495

RESULT 7
US-09-836-821-6
Sequence 6, Application US/09836821
Publication No. US20030087405A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.

TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/836.821
CURRENT FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: 09/105.537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRF
ORGANISM: Streptomyces venezuelae
US-09-836-821-6

Query Match 50.5%; Score 385.5; DB 9; Length 11877;
Best Local Similarity 54.8%; Pred. No. 1.6e-25;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YMLESTPGAGNVSAAGLDPTHEPLLGATLELATDGGALLAGRLSLRSHPMADHAGVT 60
DB 7355 YWQPDLSAAGDITTSAGLGAHEHPLGAAVALADSDCLLTGSLRTHPMADHAGVT 7414
QY 61 VLLSGATFLELALHAGTYGCDRVDELTLHAPLVVPVDGVSVOGVAADGEGRLVSV 120
DB 7415 VLLPGTAFVELAFRAADQVGCDELVEELTLDAPLVLPFRGAVRVOLSVGASDESGRRTFGL 7474
QY 121 YARGSGACGGGASGGVWTCASGVL 146
DB 7475 YAHPEADAPG-----EAEWTRHATGVL 7495

RESULT 8
US-09-861-289-6
Sequence 6, Application US/09861289
Patent No. US2002010897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/861.289
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105.537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRF
ORGANISM: Streptomyces venezuelae
US-09-861-289-6

Query Match 50.5%; Score 385.5; DB 10; Length 11877;
Best Local Similarity 54.8%; Pred. No. 1.6e-25;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YMLESTPGAGNVSAAGLDPTHEPLLGATLELATDGGALLAGRLSLRSHPMADHAGVT 60
DB 7355 YWQPDLSAAGDITTSAGLGAHEHPLGAAVALADSDCLLTGSLRTHPMADHAGVT 7414
QY 61 VLLSGATFLELALHAGTYGCDRVDELTLHAPLVVPVDGVSVOGVAADGEGRLVSV 120
DB 7415 VLLPGTAFVELAFRAADQVGCDELVEELTLDAPLVLPFRGAVRVOLSVGASDESGRRTFGL 7474
QY 121 YARGSGACGGGASGGVWTCASGVL 146
DB 7475 YAHPEADAPG-----EAEWTRHATGVL 7495

RESULT 9
US-09-988-384B-6

```
/ Sequence 6, Application US/09988384B
/ Publication No. US20030073824A1
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, D.H.
/ APPLICANT: Liu, H.
/ APPLICANT: Xue, Y.
/ APPLICANT: Zhao, L.
/ TITLE OF INVENTION: DNA encoding methymycin and pikromycin
/ FILE REFERENCE: 600.536US1
/ CURRENT APPLICATION NUMBER: US/09/988,384B
/ PRIOR FILING DATE: 2001-11-19
/ PRIOR APPLICATION NUMBER: PCT/US99/14398
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: US 09/105,537
/ PRIOR FILING DATE: 1998-06-26
/ NUMBER OF SEQ ID NOS: 53
/ SEQ ID NO 6
/ LENGTH: 12199
/ TYPE: PRT
/ ORGANISM: Streptomyces venezuelae
US-09-988-384B-6

Query Match          50.5%; Score 385.5; DB 9; Length 12199;
Best Local Similarity 54.8%; Pred. No. 1.6e-25;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YMLESTPGAGNVSAAGLDPTHEHPLGATLELATDGGALLAGRLSLRSHPLADHAGCT 60
DB 7677 YWPOPLSAGADITTSAGAAHPLLGAAVLAASDGLTGSJSLRTHPLADHAGCT 7736
QY 61 VLLSGATFLEIALHAGTYGCDRVDELTHAPLVVPVDSGVSVGVAAADGSRRLVSV 120
DB 7737 VLLGTAFLEIALHAGTYGCDRVDELTHAPLVVPVDSGVSVGVAAADGSRRLVSV 7736
QY 121 YARGGSACGGGAGGVTCHASGVL 146
DB 7797 YAHPEDAFG-----EAETRHATGVL 7817

RESULT 10
US-09-860-846-2
/ Sequence 2, Application US/09860846
/ Patent No. US20020164742A1
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, D.H.
/ APPLICANT: Liu, H.
/ APPLICANT: Xue, Y.
/ APPLICANT: Zhao, L.
/ TITLE OF INVENTION: DNA encoding methymycin and pikromycin
/ FILE REFERENCE: 600.438US1
/ CURRENT APPLICATION NUMBER: US/09/860,846
/ CURRENT FILING DATE: 2001-05-18
/ PRIOR APPLICATION NUMBER: 09/105,537
/ PRIOR FILING DATE: 1998-06-26
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2
/ LENGTH: 5215
/ TYPE: PRT
/ ORGANISM: Streptomyces venezuelae
US-09-860-846-2

Query Match          42.4%; Score 324; DB 9; Length 5215;
Best Local Similarity 52.0%; Pred. No. 2.2e-20;
Matches 77; Conservative 9; Mismatches 56; Indels 6; Gaps 4;
```

```
QY 119 SYVARGGSACGGGAGGVTCHASGVL 146
DB 2043 GYVSTPDSGDTGDDAP-REWTRHVSGL 2069

RESULT 11
US-09-988-384B-2
/ Sequence 2, Application US/09988384B
/ Publication No. US20030073824A1
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, D.H.
/ APPLICANT: Liu, H.
/ APPLICANT: Xue, Y.
/ APPLICANT: Zhao, L.
/ TITLE OF INVENTION: DNA encoding methymycin and pikromycin
/ FILE REFERENCE: 600.536US1
/ CURRENT APPLICATION NUMBER: US/09/988,384B
/ PRIOR FILING DATE: 2001-11-19
/ PRIOR APPLICATION NUMBER: PCT/US99/14398
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: US 09/105,537
/ PRIOR FILING DATE: 1998-06-26
/ NUMBER OF SEQ ID NOS: 53
/ SEQ ID NO 2
/ LENGTH: 5215
/ TYPE: PRT
/ ORGANISM: Streptomyces venezuelae
US-09-988-384B-2

Query Match          42.4%; Score 324; DB 9; Length 5215;
Best Local Similarity 52.0%; Pred. No. 2.2e-20;
Matches 77; Conservative 9; Mismatches 56; Indels 6; Gaps 4;

QY 1 YMLESTPGAGNVSAAGLDPTHEHPLGATLELATDGGALLAGRLSLRSHPLADHAGCT 60
DB 1926 YWL--TPEARTDARAAGFDPARHPLTTTVEVAGDGVLLTGRSLSTDQPLADHVNCA 1983
QY 61 VLLSGATFLEIALHAGTYGCDRVDELTHAPLVVPVDSGVSVGVAAADGSRRLVSV 118
DB 1984 VLLPATRFELEIALAAGDHVGAVEREITLLEAPLVLPERGAVRQGVV-SCDGESEPAQRTF 2042
QY 119 SYVARGGSACGGGAGGVTCHASGVL 146
DB 2043 GYVSTPDSGDTGDDAP-REWTRHVSGL 2069

RESULT 12
US-09-836-821-2
/ Sequence 2, Application US/09836821
/ Publication No. US20030087405A1
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, D.H.
/ APPLICANT: Liu, H.
/ APPLICANT: Xue, Y.
/ APPLICANT: Zhao, L.
/ TITLE OF INVENTION: DNA encoding methymycin and pikromycin
/ FILE REFERENCE: 600.438US1
/ CURRENT APPLICATION NUMBER: US/09/836,821
/ CURRENT FILING DATE: 2001-04-17
/ PRIOR APPLICATION NUMBER: 09/105,537
/ PRIOR FILING DATE: 1998-06-26
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2
/ LENGTH: 5215
/ TYPE: PRT
/ ORGANISM: Streptomyces venezuelae
US-09-836-821-2

Query Match          42.4%; Score 324; DB 9; Length 5215;
Best Local Similarity 52.0%; Pred. No. 2.2e-20;
Matches 77; Conservative 9; Mismatches 56; Indels 6; Gaps 4;
```

QY 1 YWLESTPGAGNVSAAGLDPTHEPLGATLELATDGGALLAGRLSLRSHPLADHAVGT 60
1926 YWL--TEARTDARALGFDPARHPLTTTVEVAGGDEVLLTGRSLTDQPLADHVNAGA 1983
DB 61 VLLSGATFLELALHAGTYGCDRVDELTLHAPLVVPVPGSVVQVGAADGE--GRLV 118
1984 VLLPATFLELALAAAGHVGAVRVEELTLERPLVPERGAVRIQGV--SGDGSPAGRTF 2042
QY 119 SYVARGGAGCGGAGSGVWTCASGVL 146
2043 GVYSTPDSGDTGDDAP-REWTRHVSGLV 2069
DB 2043 GVYSTPDSGDTGDDAP-REWTRHVSGLV 2069
RESULT 13
US-09-861-289-2
Sequence 2, Application US/09861289
Patent No. US20020110897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438US1
CURRENT APPLICATION NUMBER: US/09/861,289
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 5215
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-861-289-2
Query Match 42.4%; Score 324; DB 10; Length 5215;
Best Local Similarity 52.0%; Pred. No. 2,2e-20;
Matches 77; Conservative 9; Mismatches 56; Indels 6; Gaps 4;
QY 1 YWLESTPGAGNVSAAGLDPTHEPLGATLELATDGGALLAGRLSLRSHPLADHAVGT 60
1926 YWL--TEARTDARALGFDPARHPLTTTVEVAGGDEVLLTGRSLTDQPLADHVNAGA 1983
DB 61 VLLSGATFLELALHAGTYGCDRVDELTLHAPLVVPVPGSVVQVGAADGE--GRLV 118
1984 VLLPATFLELALAAAGHVGAVRVEELTLERPLVPERGAVRIQGV--SGDGSPAGRTF 2042
QY 119 SYVARGGAGCGGAGSGVWTCASGVL 146
2043 GVYSTPDSGDTGDDAP-REWTRHVSGLV 2069
DB 2043 GVYSTPDSGDTGDDAP-REWTRHVSGLV 2069
RESULT 14
US-09-793-708-1
Sequence 1, Application US/09793708
Patent No. US20030104597A1
GENERAL INFORMATION:
APPLICANT: Ashley, Gary
APPLICANT: Bertlach, Melanie C.
APPLICANT: Bertlach, Mary C.
APPLICANT: McDaniell, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002121
CURRENT APPLICATION NUMBER: US/09/793,708
CURRENT FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: US 09/657,440
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: US 09/320,878
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: US 09/141,908

PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: US 09/073,538
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: US 08/846,247
PRIOR FILING DATE: 1997-04-30
PRIOR APPLICATION NUMBER: US 60/134,990
PRIOR FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 4551
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-793-708-1
Query Match 41.9%; Score 320; DB 9; Length 4551;
Best Local Similarity 47.1%; Pred. No. 4.4e-20;
Matches 74; Conservative 19; Mismatches 46; Indels 18; Gaps 4;
QY 1 YWLESTPGAG--NVSAAGLDPTHEPLGATLELATDGGALLAGRLSLRSHPLADHAV 57
3494 YWLDA--PADTVADTRAGLGTADHPLGAVVSLDRDGLLTGRSLRTHPWLADHAV 3551
DB 58 GGTVLLSGATFLELALHAGTYGCDRVDELTLHAPLVVPVPGSVVQVGAADGE--- 113
3552 LGSVLLPGAMVLELAHAAESAGLRDVRRELTLEPLVPERHGVLEVRTVGABGEPGGE 3611
QY 114 ----GRLVSVVARGGAGCGGAGSGVWTCASGVL 146
3612 SAGDGRPVSLHSLRLDA-----PAGTAWSCATGILL 3643
DB 3612 SAGDGRPVSLHSLRLDA-----PAGTAWSCATGILL 3643
RESULT 15
US-09-860-846-31
Sequence 31, Application US/09860846
Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438US1
CURRENT APPLICATION NUMBER: US/09/860,846
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31
LENGTH: 4613
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-860-846-31
Query Match 41.9%; Score 320; DB 9; Length 4613;
Best Local Similarity 47.1%; Pred. No. 4.5e-20;
Matches 74; Conservative 19; Mismatches 46; Indels 18; Gaps 4;
QY 1 YWLESTPGAG--NVSAAGLDPTHEPLGATLELATDGGALLAGRLSLRSHPLADHAV 57
3556 YWLDA--PADTVADTRAGLGTADHPLGAVVSLDRDGLLTGRSLRTHPWLADHAV 3613
DB 58 GGTVLLSGATFLELALHAGTYGCDRVDELTLHAPLVVPVPGSVVQVGAADGE--- 113
3614 LGSVLLPGAMVLELAHAAESAGLRDVRRELTLEPLVPERHGVLEVRTVGABGEPGGE 3673
QY 114 ----GRLVSVVARGGAGCGGAGSGVWTCASGVL 146
3674 SAGDGRPVSLHSLRLDA-----PAGTAWSCATGILL 3705
DB 3674 SAGDGRPVSLHSLRLDA-----PAGTAWSCATGILL 3705

Search completed: June 17, 2003, 13:46:46

Sun Jun 22 13:12:04 2003

us-09-914-286-3_copy_2983_3128.rpb

Page 6

Job time : 5.54907 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using SW model

Run on: June 17, 2003, 12:54:32 / Search time 2.65363 seconds
(without alignments)
1618.821 Million cell updates/sec

Title: US-09-914-286-3_COPY_2983_3128

Perfect score: 764

Sequence: 1 YWLESTOPGAGNVSAAGLDP.....ACGGGASGGVWTCASGV 146

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	385.5	50.5	3739	3	US-09-320-878-2
2	385.5	50.5	3739	4	US-09-105-537-3
3	385.5	50.5	11877	4	US-09-105-537-6
4	365.5	47.8	5588	4	US-09-036-987A-6
5	365.5	47.8	5588	4	US-09-370-700-6
6	342	44.8	2152	4	US-09-036-987A-3
7	342	44.8	2152	4	US-09-370-700-3
8	333.5	43.7	4472	2	US-08-804-227C-2
9	324	42.4	5215	4	US-09-105-537-2
10	320	41.9	4551	3	US-09-320-878-1
11	320	41.9	4551	4	US-09-105-537-31
12	319	41.8	3816	4	US-09-428-517-3
13	316.5	41.4	4545	2	US-08-804-227C-14
14	316.5	41.4	4550	2	US-08-804-227C-8
15	316.5	41.4	4550	2	US-08-804-198-2
16	311.5	40.8	4928	4	US-09-036-987A-5
17	311.5	40.8	4928	4	US-09-370-700-5
18	308.5	40.4	3728	2	US-08-804-227C-4
19	291	36.3	5087	4	US-09-144-085-1
20	277.5	36.3	6095	4	US-09-144-085-2
21	276.5	36.2	3724	2	US-08-804-227C-10
22	276.5	36.2	3724	2	US-08-804-198-4
23	265	34.7	1864	2	US-08-804-227C-3
24	246.5	32.3	3567	2	US-07-642-734C-4
25	246.5	32.3	3567	3	US-08-439-009A-4
26	213	27.9	1996	2	US-08-804-227C-9
27	213	27.9	1996	2	US-08-804-198-3

28	155	20.3	1832	3	US-09-335-409-4	Sequence 4, Appl
29	155	20.3	1832	4	US-09-568-102-4	Sequence 4, Appl
30	155	20.3	1832	4	US-09-567-969-4	Sequence 4, Appl
31	155	20.3	1832	4	US-09-568-480-4	Sequence 4, Appl
32	155	20.3	1832	4	US-09-568-486-4	Sequence 4, Appl
33	155	20.3	1832	4	US-09-568-472-4	Sequence 4, Appl
34	155	20.3	1832	4	US-09-567-899-4	Sequence 4, Appl
35	155	20.3	1832	4	US-09-335-409-5	Sequence 5, Appl
36	155	20.3	1832	4	US-09-568-102-5	Sequence 5, Appl
37	155	20.3	1832	4	US-09-567-969-5	Sequence 5, Appl
38	155	20.3	1832	4	US-09-568-480-5	Sequence 5, Appl
39	155	20.3	1832	4	US-09-568-486-5	Sequence 5, Appl
40	155	20.3	1832	4	US-09-568-472-5	Sequence 5, Appl
41	155	20.3	1832	4	US-09-567-899-5	Sequence 5, Appl
42	154	20.2	1832	4	US-09-413-814-67	Sequence 67, Appl
43	154	20.2	1832	3	US-09-335-409-6	Sequence 6, Appl
44	154	20.2	1832	4	US-09-568-102-6	Sequence 6, Appl
45	154	20.2	1832	4	US-09-567-969-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-320-878-2
Sequence 2, Application US/09320878A
Patent No. 6117659
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062202120
CURRENT APPLICATION NUMBER: US/09320, 878A
EARLIER APPLICATION NUMBER: CIP OF 09/141, 908
EARLIER FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: CIP OF 09/073, 538
EARLIER FILING DATE: 1998-05-06
EARLIER APPLICATION NUMBER: CIP OF 08/846, 247
EARLIER FILING DATE: 1997-04-30
EARLIER APPLICATION NUMBER: 60/119, 139
EARLIER FILING DATE: 1999-02-08
EARLIER APPLICATION NUMBER: 60/100, 880
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: 60/087, 080
EARLIER FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 2
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-320-878-2

Query Match 50.5%; Score 385.5; DB 3; Length 3739;
Best Local Similarity 54.8%; Pred. No. 3.5e-32;

Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY	1	YWLESTOPGAGNVSAAGLDPTEHPILGATLELTGALLAGRLSRSHPMWLDHVAAGT 60
DB	2406	YWQPPISAGDITTSGLGAERHPLGAAVALADSDGCLTSLRTHWLDHVAAGT 2465
QY	61	VLLSGATFLELALHAGTYGCRVDELTLHAPLVVDGVSVOGVAAADGGRLLVSV 120
DB	2466	VLLPGAFVFLAIPRADQVCGDLVBELTLDAPLVLRRAVVRVQLSVGASDSGRRTPGL 2525
QY	121	YARGSGACGGGASGGVWTCASGV 146
DB	2526	YAHEDAPG-----EAETRHATGV 2546

RESULT 2

US-09-105-537-33
; Sequence 33, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438051
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 3739
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-33

Query Match 50.5%; Score 385.5; DB 4; Length 3739;
Best Local Similarity 54.8%; Pred. No. 3.5e-32;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YWLESTPGAGNVAAGLDPTHEPLLGATLELATDGGALLAGRLSRSHPLADHAYGVT 60
DB 2406 YWPOPDLSAAGDITSAGIGAAEHPLLGAAVALADSDGCLTGSLSLTHPLADHAYAGT 2465
QY 61 VLLSGATFLELALHAGTYVGCGRVDELTHAPLVVPDGVSVGVAAADGEGRLVSV 120
DB 2466 VLLPGTAFVLELAFRAGQVGCDELVEELTLDPVLPRGAARVQLSVGASDESGRTFTGL 2525
QY 121 YARGGSACGGGAGSGVWTCASGVL 146
DB 2526 YAHPEAPG-----EAEWTRHATGVL 2546

RESULT 3

US-09-105-537-6
; Sequence 6, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438051
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 11877
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-6

Query Match 50.5%; Score 385.5; DB 4; Length 11877;
Best Local Similarity 54.8%; Pred. No. 1.6e-31;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YWLESTPGAGNVAAGLDPTHEPLLGATLELATDGGALLAGRLSRSHPLADHAYGVT 60
DB 7355 YWPOPDLSAAGDITSAGIGAAEHPLLGAAVALADSDGCLTGSLSLTHPLADHAYAGT 7414
QY 61 VLLSGATFLELALHAGTYVGCGRVDELTHAPLVVPDGVSVGVAAADGEGRLVSV 120
DB 7415 VLLPGTAFVLELAFRAGQVGCDELVEELTLDPVLPRGAARVQLSVGASDESGRTFTGL 7474
QY 121 YARGGSACGGGAGSGVWTCASGVL 146

DB 7475 YAHPEAPG-----EAEWTRHATGVL 7495

RESULT 4

US-09-036-987A-6
; Sequence 6, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5588 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-036-987A-6

Query Match 47.8%; Score 365.5; DB 4; Length 5588;
Best Local Similarity 47.9%; Pred. No. 8.3e-30;
Matches 70; Conservative 21; Mismatches 50; Indels 5; Gaps 1;

QY 1 YWLESTPGAGNVAAGLDPTHEPLLGATLELATDGGALLAGRLSRSHPLADHAYGVT 60
DB 906 YWLESAEYDAGDGSGLSAGLAEHPLLGAAVTLADAGGFLITGSLVKTQPMADHAYGGA 965
QY 61 VLLSGATFLELALHAGTYVGCGRVDELTHAPLVVPDGVSVGVAAADGEGRLVSV 120
DB 966 VLLPGTAFVLELAFRAGQVGCDELVEELTLDPVLPRGATGAVQVQIANGPDEAGRSVRV 1025
QY 121 YARGGSACGGGAGSGVWTCASGVL 146
DB 1026 H-----SCRDAVAPDQSWTCATGTL 1046

RESULT 5

US-09-370-700-6
; Sequence 6, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:

APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patti J
APPLICANT: Turner, Jan R
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 6
LENGTH: 5588
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-6

Query Match 47.8%; Score 365.5; DB 4; Length 5588;
Best Local Similarity 47.9%; Pred. No. 8.3e-30;
Matches 70; Conservative 21; Mismatches 50; Indels 5; Gaps 1;

QY 1 YMLESTPGAGNVSAAGLDPTHEPLIGATLELATDGGALLAGRLSRHPMLADHAVGCT 60
DB 906 YMLSAEYDGDGDSVGLLSAHEPLIGAVTLADAGFLITGLSVTQWMLADHVVGGA 965

QY 61 VLLSGATFLELALHAGTVGCDRVDLTTLHAPLVVPDGVSVQVGVAAADGGRRLVSV 120
DB 966 ILLPRTAFVEMLRADQVGCDDLEELSTPLVLPATGAVQVQIAGVGPDEAGRREVRV 1025

QY 121 YARGGACGGGASGCVWTCASGVL 146
DB 1026 H-----SCRDDAVPQDSWTCATGTL 1046

RESULT 6

US-09-036-987A-3
Sequence 3, Application US/09036987A
Patent No. 6143526

GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Meilo, Donald J.
APPLICANT: Treadway, Patti J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R
REGISTRATION NUMBER: 28,479

REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2152 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-3

Query Match 44.8%; Score 342; DB 4; Length 2152;
Best Local Similarity 47.9%; Pred. No. 7.9e-28;
Matches 70; Conservative 27; Mismatches 41; Indels 8; Gaps 2;

QY 1 YMLESTPGAGNVSAAGLDPTHEPLIGATLELATDGGALLAGRLSRHPMLADHAVGCT 60
DB 920 YMLVPSD--SGDVTGAGLAGAHEPLIGAVPVAGDDEVLLTGRISVTRTHPMLAHEHVLGE 977

QY 61 VLLSGATFLELALHAGTVGCDRVDLTTLHAPLVVPDGVSVQVGVAAADGGRRLVSV 120
DB 978 VIVAGTALIEILAHAEERLGCERVELTEAPLVLPERGAIOVLVGAPENSGRPMAL 1037

QY 121 YARGGACGGGASGCVWTCASGVL 146
DB 1038 YSR-----PEGAAEHDTWRHATGRL 1057

RESULT 7

US-09-370-700-3
Sequence 3, Application US/09370700
Patent No. 6274350

GENERAL INFORMATION:
APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patti J
APPLICANT: Turner, Jan R
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 3
LENGTH: 2152
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-3

Query Match 44.8%; Score 342; DB 4; Length 2152;
Best Local Similarity 47.9%; Pred. No. 7.9e-28;
Matches 70; Conservative 27; Mismatches 41; Indels 8; Gaps 2;

QY 1 YMLESTPGAGNVSAAGLDPTHEPLIGATLELATDGGALLAGRLSRHPMLADHAVGCT 60
DB 920 YMLVPSD--SGDVTGAGLAGAHEPLIGAVPVAGDDEVLLTGRISVTRTHPMLAHEHVLGE 977

QY 61 VLLSGATFLELALHAGTVGCDRVDLTTLHAPLVVPDGVSVQVGVAAADGGRRLVSV 120
DB 978 VIVAGTALIEILAHAEERLGCERVELTEAPLVLPERGAIOVLVGAPENSGRPMAL 1037

QY 121 YARGGACGGGASGCVWTCASGVL 146
DB 1038 YSR-----PEGAAEHDTWRHATGRL 1057

RESULT 8

```

US-08-804-227C-2
; Sequence 2, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kuhose, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4472 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-804-227C-2

Query Match 43.7%; Score 333.5; DB 2; Length 4472;
Best Local Similarity 45.7%; Pred. No. 1.7e-26;
Matches 69; Conservative 21; Mismatches 56; Indels 5; Gaps 2;

QY 1 YMLESTPGAGNVA--GDDPTEHPILGATLELATDGGALLAARLSLRSHPWLTADAVG 58
DB 3490 YMAVTSFAGVGDPAAGRFMTWEDHPLRGGLPLADSGERFVAGRLAGSEHMDLTDAVS 3549
QY 59 GTVLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPVDSGVSVQVGAADGGRRLV 118
DB 3550 GTVLLPCTAFVETALHAGATGCGRLSELSEAPLVLPAGGVAVQVQKRSAAADSGRRV 3609
QY 119 SVYARGGSA--CGGGASGGGVTCHASGVL 146
DB 3610 AIHSAPPAVAHSAAGCGDSAGVWTRHGEGTL 3640

RESULT 9
US-09-105-537-2
; Sequence 2, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2

```

```

; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-105-537-2

Query Match 42.4%; Score 324; DB 4; Length 5215;
Best Local Similarity 52.0%; Pred. No. 2.1e-25;
Matches 77; Conservative 9; Mismatches 56; Indels 6; Gaps 4;

QY 1 YMLESTPGAGNVAAGLDPTBHPILGATLELATDGGALLAARLSLRSHPWLTADHAGCT 60
DB 1926 YML--TPEARTDARALGFDPARHPILTTVEVAGGQVLLTGRSLTDQPMADHWNGA 1983
QY 61 VLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPVDSGVSVQVGAADG--GRLV 118
DB 1984 VLLPATRFLALAAAGDHVAVVEITLLEAPLVLPBGRANVRIQGV--SGDGSPACRTF 2042
QY 119 SVYARGGSAAGGASGGGVTCHASGVL 146
DB 2043 GVVSTPDSGDTGDAP-REWTRHVSGL 2069

RESULT 10
US-09-320-878-1
; Sequence 1, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; EARLIER FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4551
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-320-878-1

Query Match 41.9%; Score 320; DB 3; Length 4551;
Best Local Similarity 47.1%; Pred. No. 4.8e-25;
Matches 74; Conservative 19; Mismatches 46; Indels 18; Gaps 4;

QY 1 YMLESTPGAG--NVSAGLDPTBHPILGATLELATDGGALLAARLSLRSHPWLTADHAG 57
DB 3494 YMLDA--PAADTVADTAGLGLGADHPILGAVSLDPDQGLITGRSLRTHPWLTADHAG 3551
QY 58 GTVLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPVDSGVSVQVGAADG--GRLV 113
DB 3552 LGSVLLPGAMVETLAAHAAESAGLRDVRETLLEPLVLPBHGGEVLRVTGADAGPGR 3611
QY 114 ----GRLVSVYARGGSAAGGASGGGVTCHASGVL 146
DB 3612 SAGDGPVSLHSLRLA-----PAGTAWSCATGTL 3643

```

Oy I YMLE-STOPOA -GNVSAAGLDPEHPLLGATLEATDGCALLGRLSLRSHPMADHAV 57
 Db 2471 YMKPAPAPCANLGVASVGLTAACHPLGLCAVNEHPSDGLVLTQOISLTHPEMLADHEV 2500

RESULT 14
US-08-804-227C-8
; Sequence 8, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:

```

APPLICANT: DeHoff, Bradley S.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rosteck, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 4550 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-227C-8

Query Match 41.4%; Score 316.5; DB 2; Length 4550;
Best Local Similarity 48.3%; Pred. No. 1.1e-24;
Matches 73; Conservative 14; Mismatches 55; Indels 9; Gaps 3;

QY 1 YWLESTOPGAGNVSA--GLDPTHEPLGATLELATDGGALLAGRLSLRSHPLADHAV 57
Db 3533 YMATGSVTGATGTSAARFGLWKDHPFLSGATPIAGSGALLTGRVGLAHPWLADHAI 3592

QY 58 GGTVLGSGATFLELALHAGTYVGCDDVDELTLHAPLVVPVDSVSVQVGVAAADGEGRL 117
Db 3593 SGTVLPGTAIADLLRAVEVGAGVBEVLTLEHPLLPERGGLHVOVLVEAADGGRRA 3652

QY 118 VSVYARGGSACGGGASG--GVWTCASGVL 146
Db 3653 VAVAAP--EGPGRDGEQEWTRHAEGLV 3679

RESULT 15
US-08-804-198-2
Sequence 2, Application US/08804198
Patent No. 5945320
GENERAL INFORMATION:
APPLICANT: Burgett, Stanley G.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rao, Nagara R.
APPLICANT: Richardson, Mark A.
APPLICANT: Rosteck, Paul R., Jr.
TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
```

```

MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4550 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-2

Query Match 41.4%; Score 316.5; DB 2; Length 4550;
Best Local Similarity 48.3%; Pred. No. 1.1e-24;
Matches 73; Conservative 14; Mismatches 55; Indels 9; Gaps 3;

QY 1 YWLESTOPGAGNVSA--GLDPTHEPLGATLELATDGGALLAGRLSLRSHPLADHAV 57
Db 3533 YMATGSVTGATGTSAARFGLWKDHPFLSGATPIAGSGALLTGRVGLAHPWLADHAI 3592

QY 58 GGTVLGSGATFLELALHAGTYVGCDDVDELTLHAPLVVPVDSVSVQVGVAAADGEGRL 117
Db 3593 SGTVLPGTAIADLLRAVEVGAGVBEVLTLEHPLLPERGGLHVOVLVEAADGGRRA 3652

QY 118 VSVYARGGSACGGGASG--GVWTCASGVL 146
Db 3653 VAVAAP--EGPGRDGEQEWTRHAEGLV 3679

Search completed: June 17, 2003, 13:09:30
Job time : 3.65363 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 12:50:57 ; Search time 7.15215 Seconds
(without alignments)
2720.102 Million cell updates/sec

Title: US-09-914-286-3_COPY_2983_3128

Perfect score: 764
Sequence: 1 YWLESTOPGAGNVSAAGLDP.....ACGGGAGSGVWTCASGVL 146

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID82/gcgdata/geneseq/geneeqp-emb1/AA1980.DAT.*
2: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT.*
3: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT.*
4: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT.*
5: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT.*
6: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT.*
7: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT.*
8: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT.*
9: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT.*
10: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT.*
11: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT.*
12: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT.*
13: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT.*
14: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT.*
15: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT.*
16: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT.*
17: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT.*
18: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT.*
19: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT.*
20: /SID82/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT.*
21: /SID82/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT.*
22: /SID82/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT.*
23: /SID82/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	764	100.0	3972	21 AAB23749	S. avermitilis ave
2	764	100.0	3972	22 AAG65264	Streptomyces nous
3	764	100.0	3972	22 AAG65268	Streptomyces nous
4	753	98.6	146	21 AAB23753	S. avermitilis ave
5	585.5	76.6	5532	21 AAB23752	S. avermitilis ave
6	585.5	76.6	5532	22 AAG65267	Streptomyces averm
7	385.5	50.5	3739	21 AAB18638	Amino acid sequenc
8	385.5	50.5	3739	21 AAY77193	S. venezuelae macr
9	385.5	50.5	3739	21 AAY77201	S. venezuelae pik
10	385.5	50.5	3739	21 AAY67202	Narabolide synth

11	385.5	50.5	12199	21 AAY77180	S. venezuelae pik
12	372	48.7	7068	22 AAB10142	Streptomyces nous
13	372	48.7	9477	22 AAB10144	Streptomyces nous
14	365.5	47.8	5588	20 AAY93301	Spn2 a polyketide
15	365.5	47.8	5588	22 AAB70969	S. spinosa protein
16	357	46.7	4881	21 AAB23751	S. avermitilis ave
17	357	46.7	4881	22 AAG65266	Streptomyces averm
18	354.5	46.4	11096	22 AAB10129	Streptomyces nous
19	351.5	46.0	6239	21 AAB23750	S. avermitilis ave
20	351.5	46.0	6239	22 AAG65265	Streptomyces averm
21	348	45.5	6797	22 AAB31558	Pimaricin biosynth
22	342	44.8	2152	20 AAY9298	Spn2 a polyketide
23	342	44.8	2152	22 AAB70966	S. spinosa protein
24	333.5	43.7	4472	18 AAB22601	Tylectone synthase
25	332	43.5	5069	19 AAW52846	A. mediterranei r1
26	324	42.4	4630	18 AAB19629	Streptomyces venez
27	324	42.4	4630	21 AAY77177	S. venezuelae vep
28	320	41.9	4551	21 AAB18637	Amino acid sequenc
29	320	41.9	4551	21 AAY77201	Narabolide synth
30	320	41.9	4613	21 AAY77192	S. venezuelae macr
31	320	41.9	4613	21 AAY77200	S. venezuelae pik
32	319	41.8	3816	21 AAY92708	S. antibioticus 8
33	316.5	41.4	4545	18 AAB22611	Hybrid ermC/tylG O
34	316.5	41.4	4550	18 AAB23716	Platenolide synth
35	316.5	41.4	4550	18 AAB22606	Platenolide synth
36	311.5	40.8	4924	22 AAB70968	S. spinosa protein
37	311.5	40.8	4928	20 AAY93900	Spn2 a polyketide
38	309.5	40.5	5435	22 AAB10145	Streptomyces nous
39	308.5	40.4	1891	19 AAB52844	Amycolatopsis medl
40	308.5	40.4	3413	19 AAW52849	A. mediterranei r1
41	308.5	40.4	3729	18 AAB22603	Tylectone synthase
42	304	39.8	1615	22 AAG81211	Mycobacterium tub
43	300	39.3	1721	19 AAW52847	A. mediterranei r1
44	291	38.1	5087	22 AAB10700	Amino acid sequenc
45	279.5	36.6	1366	22 AAB10127	Streptomyces nous

ALIGNMENTS

RESULT 1
AAB23749 standard; Protein; 3972 AA.
XX
AC AAB23749;
XX
DT 10-JAN-2001 (first entry)
XX
DE S. avermitilis avermectin aglycon synthase protein SEQ ID NO:3.
XX
KW Streptomyces avermitilis; avermectin aglycon synthase; biosynthetis;
KW multifunctional enzyme; polyketide; avermectin; veterinary drug;
KW agrochemical.
XX
OS Streptomyces avermitilis.
XX
PN MO200050605-A1.
XX
PD 31-AUG-2000.
XX
PF 23-FEB-2000; 2000WO-JP01041.
XX
PR 24-FEB-1999; 99JP-0046961.
XX
PA (KITA) KITASATO INST.
XX
PI Omura S, Ikeda H;
XX
DR WPI; 2000-565458/52.
XX
DR N-PSDB; AAA92301.
XX
PT Avermectin aglycone synthase DNA and proteins encoded by all or part of
it for the production of avermectin and its derivatives for drug and

PT agrochemical use -
XX
PS Claim 32; Page 203-223; 314pp; Japanese.
XX
CC The present sequence represents an avermectin aglycon synthase protein.
CC Also described are: (1) polypeptides encoded by all or part of the DNA;
CC (2) expression vectors containing the DNA; (3) host cells transformed by
CC the vectors; (4) preparation of the polypeptides by culture of the
CC transformants; (5) preparation of avermectin aglycon or its derivatives
CC by culture of transformed avermectin-producing microorganisms; and (6)
CC oligonucleotides of 5-60 bases in length containing sense or antisense
CC sequences from the avermectin aglycon synthase DNA. The enzymes are
CC useful for the production of modified forms of avermectin and of the
CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
CC and agrochemicals.
SQ Sequence 3972 AA:
Query Match 100.0%; Score 764; DB 21; Length 3972;
Best Local Similarity 100.0%; Pred. No. 2e-68;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YMLESTPGAGNVSAAAGLDPTHEPLTGATLELTDGALLAGRLSRSHPWLDHAYGCT 60
DB 2983 YMLESTPGAGNVSAAAGLDPTHEPLTGATLELTDGALLAGRLSRSHPWLDHAYGCT 3042
QY 61 VLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGGVSVOGVAAADGEGRLVSV 120
DB 3043 VLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGGVSVOGVAAADGEGRLVSV 3102
QY 121 YARGGSACGGGASGGVWTCASGVL 146
DB 3103 YARGGSACGGGASGGVWTCASGVL 3128
RESULT 2
AAG65264
ID AAG65264 standard; Protein; 3972 AA.
XX
AC AAG65264;
XX
DT 04-DEC-2001 (first entry)
XX
DE Streptomycetes avermilitilis protein SEQ ID NO: 4.
XX
KM Avermectin aglycone synthase; AAS; avermectin derivative;
KM drug production; veterinary drug; pesticide.
XX
OS Streptomycetes avermilitilis.
XX
PN WO200162939-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-JP01381.
XX
PR 24-FEB-2000; 2000JP-0047405.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PA (KITA) KITASATO INST.
XX
PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
XX
DR WPI; 2001-582053/65.
DR N-PSDB; AAH79277.
XX
PT New modified avermectin aglycone synthase derived from Streptomycetes
PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
PT drugs and pesticides -
XX
PS Claim 4; Page 167-180; 257pp; Japanese.
XX
CC The present invention relates to the production of modified derivatives

CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
CC avermilitilis. The activity of an acyl carrier protein (ACP),
CC beta-ketoadyl carrier protein synthase (KS), acyltransferase (AT),
CC beta-ketoadyl carrier protein reductase (KR), dehydratase (DH), enoyl
CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC suppressed. The process can be used in the production of drugs, veterinary
CC drugs and pesticides. The present sequence is an S. avermilitilis protein.
SQ Sequence 3972 AA:
Query Match 100.0%; Score 764; DB 22; Length 3972;
Best Local Similarity 100.0%; Pred. No. 2e-68;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YMLESTPGAGNVSAAAGLDPTHEPLTGATLELTDGALLAGRLSRSHPWLDHAYGCT 60
DB 2983 YMLESTPGAGNVSAAAGLDPTHEPLTGATLELTDGALLAGRLSRSHPWLDHAYGCT 3042
QY 61 VLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGGVSVOGVAAADGEGRLVSV 120
DB 3043 VLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGGVSVOGVAAADGEGRLVSV 3102
QY 121 YARGGSACGGGASGGVWTCASGVL 146
DB 3103 YARGGSACGGGASGGVWTCASGVL 3128
RESULT 3
AAG65268
ID AAG65268 standard; Protein; 3972 AA.
XX
AC AAG65268;
XX
DT 04-DEC-2001 (first entry)
XX
DE Streptomycetes avermilitilis protein derivative SEQ ID NO: 8.
XX
KM Avermectin aglycone synthase; AAS; avermectin derivative;
KM drug production; veterinary drug; pesticide.
XX
OS Synthetic.
XX
PN WO200162939-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-JP01381.
XX
PR 24-FEB-2000; 2000JP-0047405.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
PA (KITA) KITASATO INST.
XX
PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
XX
DR WPI; 2001-582053/65.
DR N-PSDB; AAH79279.
XX
PT New modified avermectin aglycone synthase derived from Streptomycetes
PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
PT drugs and pesticides -
XX
PS Claim 5; Page 235-248; 257pp; Japanese.
XX
CC The present invention relates to the production of modified derivatives
CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
CC avermilitilis. The activity of an acyl carrier protein (ACP),
CC beta-ketoadyl carrier protein synthase (KS), acyltransferase (AT),
CC beta-ketoadyl carrier protein reductase (KR), dehydratase (DH), enoyl
CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC suppressed. The process can be used in the production of drugs, veterinary
CC drugs and pesticides. The present sequence is derived from an S.
CC avermilitilis protein.

XX Sequence 3972 AA;
SQ
Query Match 100.0%; Score 764; DB 22; Length 3972;
Best Local Similarity 100.0%; Pred. No. 2e-68;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YWLESTPGAGNVSAAGLDPTHEPLTGATLELATDGGALLAGRLSRSHPWLDHAYGVT 60
DB 2983 YWLESTPGAGNVSAAGLDPTHEPLTGATLELATDGGALLAGRLSRSHPWLDHAYGVT 3042
QY 61 VLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGVSVQVGVAAADGEGRLVSV 120
DB 3043 VLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGVSVQVGVAAADGEGRLVSV 3102
QY 121 YARGSGACGGGGASGGVWTCCHASGVL 146
DB 3103 YARGSGACGGGGASGGVWTCCHASGVL 3128

RESULT 4
AAB23753
ID AAB23753 standard; Protein; 146 AA.
XX
AC AAB23753;
XX
DT 10-JAN-2001 (first entry)
XX
DE S. avermectilis avermectin aglycon synthase protein SEQ ID NO:8.
XX
KM Streptomyces avermectilis; avermectin aglycon synthase; bioynthesis;
KM multifunctional enzyme; polyketide; avermectin; veterinary drug;
KM agrochemical.
XX
OS Streptomyces avermectilis.
XX
PN WO200050605-A1.
XX
PD 31-AUG-2000.
XX
PF 23-FEB-2000; 2000WO-JP01041.
XX
PR 24-FEB-1999; 99JP-0046961.
XX
PA (KITA) KITASATO INST.
XX
PI Omura S, Ikeda H;
PI; 2000-565458/52.
DR N-PSDB; AAA92303.
XX
PT Avermectin aglycone synthase DNA and proteins encoded by all or part of
PT it for the production of avermectin and its derivatives for drug and
PT agrochemical use -
XX
PS Disclosure; Page 307-308; 314pp; Japanese.
XX
CC The present sequence represents an avermectin aglycon synthase protein.
CC Also described are: (1) polypeptides encoded by all or part of the DNA;
CC (2) expression vectors containing the DNA; (3) host cells transformed by
CC the vectors; (4) preparation of the polypeptides by culture of the
CC transformants; (5) preparation of avermectin aglycon or its derivatives
CC by culture of transformed avermectin-producing microorganisms; and (6)
CC oligonucleotides of 5-60 bases in length containing sense or antisense
CC sequences from the avermectin aglycon synthase DNA. The enzymes are
CC useful for the production of modified forms of avermectin and of the
CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
CC and agrochemicals.
XX
SQ Sequence 146 AA;
Query Match 98.6%; Score 753; DB 21; Length 146;
Best Local Similarity 98.6%; Pred. No. 4.4e-69;

Matches 144; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 YWLESTPGAGNVSAAGLDPTHEPLTGATLELATDGGALLAGRLSRSHPWLDHAYGVT 60
DB 1 YWLESTPGAGNVSAAGLDPTHEPLTGATLELATDGGALLAGRLSRSHPWLDHAYGVT 60
QY 61 VLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGVSVQVGVAAADGEGRLVSV 120
DB 61 VLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPDGVSVQVGVAAADGEGRLVSV 120
QY 121 YARGSGACGGGGASGGVWTCCHASGVL 146
DB 121 YARGSGACGGGGASGGVWTCCHASGVL 146

RESULT 5
AAB23752
ID AAB23752 standard; Protein; 5532 AA.
XX
AC AAB23752;
XX
DT 10-JAN-2001 (first entry)
XX
DE S. avermectilis avermectin aglycon synthase protein SEQ ID NO:6.
XX
KM Streptomyces avermectilis; avermectin aglycon synthase; bioynthesis;
KM multifunctional enzyme; polyketide; avermectin; veterinary drug;
KM agrochemical.
XX
OS Streptomyces avermectilis.
XX
PN WO200050605-A1.
XX
PD 31-AUG-2000.
XX
PF 23-FEB-2000; 2000WO-JP01041.
XX
PR 24-FEB-1999; 99JP-0046961.
XX
PA (KITA) KITASATO INST.
XX
PI Omura S, Ikeda H;
PI; 2000-565458/52.
DR N-PSDB; AAA92302.
XX
PT Avermectin aglycone synthase DNA and proteins encoded by all or part of
PT it for the production of avermectin and its derivatives for drug and
PT agrochemical use -
XX
PS Claim 32; Page 278-305; 314pp; Japanese.
XX
CC The present sequence represents an avermectin aglycon synthase protein.
CC Also described are: (1) polypeptides encoded by all or part of the DNA;
CC (2) expression vectors containing the DNA; (3) host cells transformed by
CC the vectors; (4) preparation of the polypeptides by culture of the
CC transformants; (5) preparation of avermectin aglycon or its derivatives
CC by culture of transformed avermectin-producing microorganisms; and (6)
CC oligonucleotides of 5-60 bases in length containing sense or antisense
CC sequences from the avermectin aglycon synthase DNA. The enzymes are
CC useful for the production of modified forms of avermectin and of the
CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
CC and agrochemicals.
XX
SQ Sequence 5532 AA;
Query Match 76.6%; Score 585.5; DB 21; Length 5532;
Best Local Similarity 76.0%; Pred. No. 5.1e-50;
Matches 111; Conservative 15; Mismatches 19; Indels 1; Gaps 1;
QY 1 YWLESTPGAGNVSAAGLDPTHEPLTGATLELATDGGALLAGRLSRSHPWLDHAYGVT 60
DB 955 YWLESTPGAGNVSAAGLDPTHEPLTGATLELATDGGALLAGRLSRSHPWLDHAYGVT 1013

QY 61 VLLSGATFLELALHAGTYVGCGRVDELTLHAPLVVVDGVSVOVGAADGEGRLVSV 120
DB 1014 VLLSGSAFVELAVGVGERTRIEQLTLVHAPLVVPGGVSVOVGAADGEGRLVSV 1073
QY 121 YARGGSACGGGASGSGVWTCHASGVL 146
DB 1074 YARGGSACGGGASGSGVWTCHASGVL 1099

RESULT 6
AAG65267
ID AAG65267 standard; Protein; 5532 AA.
XX
AC AAG65267;
XX
DT 04-DEC-2001 (first entry)
XX
DE Streptomycetes avermiltillis protein SEQ ID NO: 7.
XX
KM Avermectin aglycone synthase; AAS; avermectin derivative;
KM drug production; veterinary drug; pesticide.
XX
OS Streptomycetes avermiltillis.
XX
PN WC200162939-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-JP01381.
XX
PR 24-FEB-2000; 2000JP-0047405.
XX
PA (KYOWA) KYOWA HAKKO KOGYO KK.
PA (KITTA) KITTA INST.
XX
PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
XX
DR WPI; 2001-582053/65.
DR N-PSDB; AAH79278.
XX
PT New modified avermectin aglycone synthase derived from Streptomycetes
PT avermectillis used in production of 22,23-dihydroavermectin B1a used in
PT drugs and pesticides -
XX
PS Claim 4; Page 217-235; 257pp; Japanese.
XX
CC The present invention relates to the production of modified derivatives
CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
CC avermiltillis. The activity of an acyl carrier protein (ACP),
CC beta-ketolactoyl carrier protein synthase (KS), acyltransferase (AT),
CC beta-ketolactoyl carrier protein reductase (KR), dehydratase (DH), enoyl
CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC suppressed. The process can be used in the production of drugs, veterinary
CC drugs and pesticides. The present sequence is an S. avermiltillis protein.
XX
SQ Sequence 5532 AA;
Query Match 76.6%; Score 585.5; DB 22; Length 5532;
Best Local Similarity 76.0%; Pred. No. 5.1e-50;
Matches 111; Conservative 15; Mismatches 19; Indels 1; Gaps 1;

QY 1 YMLESTQPGAGNSAAGLPTEHPHLLGATLELATDGAALLAGRLSLRSHPLADHAGVT 60
DB 955 YMLDA-PTGAGDVAAAGLEPHEPLLAIVQALDTDCCLTLGRSLSHPLGIVEVGA 1013
QY 61 VLLSGATFLELALHAGTYVGCGRVDELTLHAPLVVVDGVSVOVGAADGEGRLVSV 120
DB 1014 VLLSGSAFVELAVGVGERTRIEQLTLVHAPLVVPGGVSVOVGAADGEGRLVSV 1073
QY 121 YARGGSACGGGASGSGVWTCHASGVL 146
DB 1074 YARGGSACGGGASGSGVWTCHASGVL 1099

RESULT 7
AAB18638
ID AAB18638 standard; Protein; 3739 AA.
XX
AC AAB18638;
XX
DT 22-JAN-2001 (first entry)
XX
DE Amino acid sequence of narbonolide synthase subunit 2 (PICAI1).
XX
KM Narbonolide synthase; polyketide synthase gene; narbonolide polyketide;
KM antibiotic; C12-hydroxylase; pick; desosamine biosynthesis;
KM desosaminyl transferase enzyme; ketolide; beta-glucosidase enzyme;
KM picromycin biosynthesis.
XX
OS Streptomycetes venezuelae.
XX
PN US6117659-A.
XX
PD 12-SEP-2000.
XX
PF 27-MAY-1999; 99US-0320878.
XX
PR 28-MAY-1998; 98US-0087080.
PR 22-SEP-1998; 98US-0100880.
PR 08-FEB-1999; 99US-0119139.
PR 20-MAY-1999; 99US-0134990.
PR 30-APR-1997; 97US-0846247.
PR 06-MAY-1998; 98US-0073538.
PR 28-AUG-1998; 98US-0141908.
XX
PA (KOSA-) KOSAN BIOSCIENCES INC.
XX
PI Ashley G, Betlach MC, Betlach M, Tang L, McDaniel R;
XX
DR WPI; 2000-610844/58.
XX
PT New recombinant pick hydroxylase gene of Streptomycetes venezuelae useful
PT for converting ketolides to antibiotics and as antibiotics and
PT intermediates in the synthesis of compounds with pharmaceutical value
XX
PS Disclosure; Columns 11-12; 117pp; English.
XX
CC The present sequence represents a narbonolide synthase subunit 2
CC (PICAI1). The nucleotide sequence encoding it is used in the course of
CC the invention. The specification describes a recombinant DNA compound
CC expressing recombinant polyketide synthase genes in host cells for the
CC production of narbonolide, narbonolide derivatives and polyketides that
CC are useful as antibiotics and as intermediates in the synthesis of
CC compounds with pharmaceutical value. The DNA compounds may also encode
CC a C12-hydroxylase (pick), desosamine biosynthesis and desosaminyl
CC transferase enzymes (useful for conversion of ketolides to antibiotics),
CC and the beta-glucosidase enzyme (involved in picromycin biosynthesis).
CC These compounds are also useful for increasing the antibiotic activity
CC of a compound relative to the unhydroxylated compound. The recombinant
CC host cells are useful as genetic systems that allow rapid engineering
CC of the narbonolide polyketide synthase. These would be valuable for
CC creating novel ketolide analogs for pharmaceutical applications.
XX
SQ Sequence 3739 AA;
Query Match 50.5%; Score 385.5; DB 21; Length 3739;
Best Local Similarity 54.8%; Pred. No. 8.3e-30;
Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YMLESTQPGAGNSAAGLPTEHPHLLGATLELATDGAALLAGRLSLRSHPLADHAGVT 60
DB 2406 YWQPDLSAAGDITTSAGLAAEPHLLGAVALADSGCLLTGSLSLRTHPWLADHAGVT 2465
QY 61 VLLSGATFLELALHAGTYVGCGRVDELTLHAPLVVVDGVSVOVGAADGEGRLVSV 120

Db 2466 VLLPGTAFVELAFRAGDQVCDLVEBELTLDAPLVLRGAVRVQLSVGASDESGRRTFGL 2525
 QY 121 YARGSGACGGGASGGVWTCASGVL 146
 Db 2526 YAHPEDAFG-----EAEWTRHATGVL 2546

RESULT 8
 AAY77193
 ID AAY77193 standard; Protein, 3739 AA.
 AC AAY77193;
 DT 05-JUN-2000 (first entry)
 XX

DE S. venezuelae macrolide biosynthetic enzyme pikAII, SEQ ID NO:33.
 KW Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
 KW neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
 KW biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;
 KW chronic obstructive pulmonary disease; respiratory inflammation;
 KW hypercholesterolaemia; crop protection agent.
 OS Streptomyces venezuelae ATCC15439.
 XX
 XX MO200000620-A2.
 PN 06-JAN-2000.
 XX
 XX 25-JUN-1999; 99MO-US14398.
 XX
 XX 26-JUN-1998; 98US-0105537.
 XX
 PA (MINU) UNIV MINNESOTA.
 PI Sherman DH, Liu H, Xue Y, Zhao L;
 DR WPI; 2000-160679/14.
 DR N-PSDB; AAZ87298.
 XX
 PT Desosamine and macrolide biosynthetic gene clusters, useful for, e .g.
 PT synthesis of methymycin and pikromycin -
 XX
 XX Claim 19; Page 403-415; 438pp; English.

The invention relates to an isolated and purified nucleic acid segment comprising a desosamine biosynthetic gene cluster, a fragment or its biologically active variant, where the nucleic acid sequence is not derived from the eryC gene cluster of *Saccharopolyspora erythraea* or *Streptomyces antibioticus*. The invention also relates to a macrolide biosynthetic gene cluster, or fragments thereof. The macrolide biosynthetic gene cluster encodes proteins which synthesise methymycin, pikromycin, neomethymycin, narbomycin or a combination of these compounds. Recombinant or augmented cells comprising the desosamine and/or macrolide biosynthetic gene clusters are useful for the production of biologically active macrolides. The macrolide biosynthetic proteins are useful for synthesis of methymycin, pikromycin, neomethymycin and narbomycin. The alternative termination of polyketide synthesis may be useful to prepare novel antibiotics and

CC AAZ87295-287302.
 XX
 SQ Sequence 3739 AA;
 Query Match 50.5%; Score 385.5; DB 21; Length 3739;
 Best Local Similarity 54.8%; Pred. No. 8.3e-30;
 Matches 80; Conservative 14; Mismatches 47; Indels 5; Gaps 1;

QY 1 YWLESTPGAGVNSAAGLDPTHEPLTGATLELATDGGALLAGRLSHPMLADHAYGCT 60
 Db 2406 YWPPDLSAAGDITTSAGLGAHEHPLGAAVALADSDCLTGSLSLTHPLADHAYAGT 2465
 QY 61 VLLSGATFLELALHAGTYVCDDRVDLTLAPLVVPVDSGVSVQVGAADSEGRRLVSV 120
 Db 2466 VLLPGTAFVELAFRAGDQVCDLVEBELTLDAPLVLRGAVRVQLSVGASDESGRRTFGL 2525
 QY 121 YARGSGACGGGASGGVWTCASGVL 146
 Db 2526 YAHPEDAFG-----EAEWTRHATGVL 2546

RESULT 9
 AAY77201
 ID AAY77201 standard; Protein, 3739 AA.
 AC AAY77201;
 DT 05-JUN-2000 (first entry)
 XX

DE S. venezuelae pik (macrolide biosynthesis) gene cluster protein #2.
 KW Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
 KW neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
 KW biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;
 KW chronic obstructive pulmonary disease; respiratory inflammation;
 KW hypercholesterolaemia; crop protection agent.
 OS Streptomyces venezuelae ATCC15439.
 XX
 XX MO200000620-A2.
 PN 06-JAN-2000.
 XX
 XX 25-JUN-1999; 99MO-US14398.
 XX
 XX 26-JUN-1998; 98US-0105537.
 XX
 PA (MINU) UNIV MINNESOTA.
 PI Sherman DH, Liu H, Xue Y, Zhao L;
 DR WPI; 2000-160679/14.
 DR N-PSDB; AAZ87318.
 XX
 PT Desosamine and macrolide biosynthetic gene clusters, useful for, e .g.
 PT synthesis of methymycin and pikromycin -
 XX
 XX Disclosure, Figure 31; 438pp; English.

The invention relates to an isolated and purified nucleic acid segment comprising a desosamine biosynthetic gene cluster, a fragment or its biologically active variant, where the nucleic acid sequence is not derived from the eryC gene cluster of *Saccharopolyspora erythraea* or *Streptomyces antibioticus*. The invention also relates to a macrolide biosynthetic gene cluster, or fragments thereof. The macrolide biosynthetic gene cluster encodes proteins which synthesise methymycin, pikromycin, neomethymycin, narbomycin or a combination of these compounds. Recombinant or augmented cells comprising the desosamine and/or macrolide biosynthetic gene clusters are useful for the production of biologically active macrolides. The macrolide biosynthetic proteins are useful for synthesis of methymycin, pikromycin, neomethymycin and narbomycin. The alternative termination of polyketide synthesis may be useful to prepare novel antibiotics and

PA (SNTF) SINTER STIPELSEN IND TEK FORSK.
 PA (ALPH-) ALPHARMA AS.
 PA (SINV-) SINVENT AS.
 PA (DZIE/) DZIEGLEWSKA H.
 PA (ZOTC/) ZOTCHEV S B.
 PA (SEKU/) SEKUROVA O N.
 PA (FJAE/) FJAEVIRIK E.
 PA (BRAU/) BRAUTASET T.
 PA (STRO/) STROM A R.
 XX
 PI Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;
 PI Valia S, Ellingsen TE, Sletta H, Gulliksen O;
 XX
 DR WPI; 2001-557614/62.
 DR N-PSDB; AAD17185.
 XX
 PT New nystatin polyketide synthase polynucleotides and polypeptides,
 PT useful as antibiotics and antifungals -
 XX
 PS Claim 15; Page 181-185; 266pp; English.
 XX
 CC The present invention relates to the cloning and sequencing of the gene
 CC cluster encoding a modular type I polyketide synthase (PKS) enzyme
 CC involved in the biosynthesis of the macrolide antibiotic nystatin.
 CC The nystatin PKS is useful as antifungal antibiotics. The present
 CC sequence is a PKS type I encoding Streptomyces noursei
 CC nystatin gene, NysI (partial).
 CC
 XX
 SQ Sequence 7068 AA;
 Query Match 48.7%; Score 372; DB 22; Length 7068;
 Best Local Similarity 51.4%; Pred. No. 4.4e-28;
 Matches 75; Conservative 18; Mismatches 47; Indels 6; Gaps 1;
 QY 1 YMESTPGAGNVAAGIDPTEHPLGATLELADGALLAGRLSHPMLADHANGCT 60
 Db 914 YMPKALOSGTADLRSGVGAANHPLLSAAVSLADAGGTLTGRLSRQTHPLADHTVRCGT 973
 QY 61 VLASGATFLALAGTGVGCDRVDELTLHAPLVVPDGVSVGVGAADGSGRLVSV 120
 Db 974 TLPGTFLPLAVAGDEVGCDRVDELTLHAPLVLPDGGVQVQVLMGNPDVSGRRIVNV 1033
 QY 121 YARGSGACGGGAGGAVTCHASGVL 146
 Db 1034 HARPD-----GDDTPTMAHATGVL 1053
 RESULT 13
 AAE10144
 ID AAE10144 standard; Protein; 9477 AA.
 XX
 AC AAE10144;
 XX
 DT 29-NOV-2001 (first entry)
 XX
 DE Streptomyces noursei nystatin gene, NysI (complete).
 XX
 KW Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;
 KW antifungal; antibiotic; PKS type I.
 XX
 OS Streptomyces noursei.
 XX
 FH Key
 FT Location/Qualifiers
 FT 41..464
 FT /label= KS15_domain
 FT /note= "ketosynthase (KS) domain"
 FT 578..889
 FT /label= ART5_domain
 FT /note= "acyltransferase (AT) domain"
 FT 903..1102
 FT /label= DH15_domain
 FT /note= "Dehydratase (DH) domain"
 FT 1446..1731
 FT Domain

FT /label= ER15_domain
 FT /note= "Enoylreductase (ER) domain"
 FT 1740..1988
 FT /label= KR15_domain
 FT /note= "ketoreductase (KR) domain"
 FT 2023..2096
 FT /label= ACP15_domain
 FT /note= "acyl carrier protein (ACP) domain"
 FT 2117..2538
 FT /label= KS16_domain
 FT /note= "ketosynthase (KS) domain"
 FT 2635..2953
 FT /label= AT16_domain
 FT /note= "acyltransferase (AT) domain"
 FT 2967..3167
 FT /label= DH16_domain
 FT /note= "Dehydratase (DH) domain"
 FT 3257..3500
 FT /label= KR16_domain
 FT /note= "ketoreductase (KR) domain"
 FT 3539..3612
 FT /label= ACP16_domain
 FT /note= "acyl carrier protein (ACP) domain"
 FT 3634..4057
 FT /label= KS17_domain
 FT /note= "ketosynthase (KS) domain"
 FT 4153..4472
 FT /label= ART17_domain
 FT /note= "acyltransferase (AT) domain"
 FT 4486..4725
 FT /label= DH17_domain
 FT /note= "Dehydratase (DH) domain"
 FT 4997..5245
 FT /label= KR17_domain
 FT /note= "ketoreductase (KR) domain"
 FT 5277..5350
 FT /label= ACP17_domain
 FT /note= "acyl carrier protein (ACP) domain"
 XX
 PN W0200159126-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 08-FEB-2001; 2001WO-GB00509.
 XX
 PR 08-FEB-2000; 2000GB-0002840.
 PR 10-APR-2000; 2000GB-0008786.
 PR 14-APR-2000; 2000GB-0005387.
 XX
 PA (UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPSELIGE.
 PA (SNTF) SINTER STIPELSEN IND TEK FORSK.
 PA (ALPH-) ALPHARMA AS.
 PA (SINV-) SINVENT AS.
 PA (DZIE/) DZIEGLEWSKA H.
 PA (ZOTC/) ZOTCHEV S B.
 PA (SEKU/) SEKUROVA O N.
 PA (FJAE/) FJAEVIRIK E.
 PA (BRAU/) BRAUTASET T.
 PA (STRO/) STROM A R.
 XX
 XX
 PI Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;
 PI Valia S, Ellingsen TE, Sletta H, Gulliksen O;
 XX
 DR WPI; 2001-557614/62.
 DR N-PSDB; AAD17186.
 XX
 PT New nystatin polyketide synthase polynucleotides and polypeptides,
 PT useful as antibiotics and antifungals -
 XX
 PS Claim 15; Page 255-260; 266pp; English.
 XX
 CC The present invention relates to the cloning and sequencing of the gene
 CC cluster encoding a modular type I polyketide synthase (PKS) enzyme

CC several domains with specific functions. The initiator module has an
 CC acyl transferase (AT) domain, and an acyl carrier protein (ACP) domain.
 CC The extender modules have the same domains plus a beta-ketothiolase (KS)
 CC domain and optionally a beta-ketoreductase domain, a dehydratase (DH)
 CC domain, and an enoyl reductase (ER) domain. The last extender module
 CC terminates with a thioester domain. The products of the genes present
 CC in the upstream region the PKS genes have been assigned names spnF-spnS
 CC AA39302-Y39315 and are responsible for different modifications in
 CC spinosyn biosynthesis. There are also two ORFs ORF15 and ORF16 present
 CC immediately upstream of spnS, producing polypeptides AA39316-Y39317,
 CC and two ORFs ORF1 and ORF2 present downstream of the PKS region
 CC producing polypeptides AA39318-Y39319. The genes are useful to improve
 CC yields of spinosyns, and for creating new spinosyns e.g. by mutagenesis,
 CC or interruption of steps in spinosyn biosynthesis. The modified
 CC spinosyns may be a new insect control agent or serve as substrates for
 CC further chemical modification and the creation of new semi-synthetic
 CC spinosyns. The genes are also useful to isolate similar sequences from S.
 CC spinosa or other species by hybridization.

XX
 SQ Sequence 5588 AA;

Query Match 47.8%; Score 365.5; DB 20; Length 5588;
 Best Local Similarity 47.9%; Pred. No. 1.5e-27;
 Matches 70; Conservative 21; Mismatches 50; Indels 5; Gaps 1;

OY 1 YMLESTOPGAGNVAAGLDPTHEPLTGATLELATDGGALLAGRLSLRSHPLADHAVGT 60
 DB 906 YMLESAEYDAGDGLSGVLSAEHPILGAANTLADAGFLTGKLSVKTQPMADHVVGGA 965
 OY 61 VLLSGATFLELALHAGTYVGCGRVDELTTHAPLVVPVDSVSVQVAAADGEGRLVSV 120
 DB 966 ILLEGTAFVEMILRAADQVGCGLIEELSLTTPVLVLPATGAVQVQIAGVGPDEAGRRSVRV 1025
 OY 121 YARGGSAACGGGAGSGVWTCASGV 146
 DB 1026 H-----SCRDDAVPQDSWTCATGTL 1046

RESULT 15

AAAB70969
 ID AAB70969 standard; Protein; 5588 AA.

XX
 AC AAB70969;

DT 28-AUG-2001 (first entry)

DE S. spinosa protein fragment encoded by ORF22, SEQ ID 50.

XX Forosamine; trimethylrhinamose; polyketide synthase; biosynthesis;
 KM spinosyn; polyketide aglycone; transgenic plant; insect resistance;
 KM macrocyclic; insecticidal; polyketide synthase.

XX Saccharopolyspora spinosa.

XX DE19957268-A1.

XX 08-MAR-2001.

XX 29-NOV-1999; 99DE-1057268.

XX 27-AUG-1999; 99DE-1040596.

XX (FARB) BAYER AG.

XX Eberz G, Moehrle V, Froede R, Velten R, Salas JA;

XX WPI; 2001-267102/28.

XX N-PSDB; AAF88339.

PT New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
 PT recombinant production of insecticidal spinosyns and their derivatives
 XX

PS Claim 57; Page 313-335; 354PD; German.

XX This invention describes a novel method nucleic acid (I) and its encoded
 CC polypeptide (II) containing at least one region that encodes an enzymatic
 CC activity involved in biosynthesis of spinosyns. (I) are used (1) to
 CC identify, inactivate or modulate genes involved in the biosynthesis of
 CC (II); (11) to generate a library of polyketide synthases; (111) for
 CC adding forosamine or trimethylrhinamose to a spinosyn or polyketide
 CC aglycone; and (1v) for recombinant production of the corresponding
 CC enzymes, which are used for production of (II), their precursors or
 CC derivatives, including production of transgenic plants that express (II)
 CC and thus have increased resistance to insects. (I) are also useful as
 CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
 CC macrocyclic with insecticidal, but not antibacterial, activity, and can
 CC also be used to raise specific antibodies, useful for identifying
 CC expression clones in a gene bank. Cells transformed with (I) may produce
 CC (II) at significantly increased levels or produce new derivatives of
 CC (II). This sequence represents an S. spinosa polyketide synthase.

XX
 SQ Sequence 5588 AA;

Query Match 47.8%; Score 365.5; DB 22; Length 5588;
 Best Local Similarity 47.9%; Pred. No. 1.5e-27;
 Matches 70; Conservative 21; Mismatches 50; Indels 5; Gaps 1;

OY 1 YMLESTOPGAGNVAAGLDPTHEPLTGATLELATDGGALLAGRLSLRSHPLADHAVGT 60
 DB 906 YMLESAEYDAGDGLSGVLSAEHPILGAANTLADAGFLTGKLSVKTQPMADHVVGGA 965
 OY 61 VLLSGATFLELALHAGTYVGCGRVDELTTHAPLVVPVDSVSVQVAAADGEGRLVSV 120
 DB 966 ILLEGTAFVEMILRAADQVGCGLIEELSLTTPVLVLPATGAVQVQIAGVGPDEAGRRSVRV 1025
 OY 121 YARGGSAACGGGAGSGVWTCASGV 146
 DB 1026 H-----SCRDDAVPQDSWTCATGTL 1046

Search completed: June 17, 2003, 13:02:33
 Job time : 8.15215 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:53:02 ; Search time 19.8736 Seconds
(without alignments)
4458.180 Million cell updates/sec

Title: US-09-914-286-3_COPY_2086_2515

Perfect score: 2265
Sequence: 1 EPIAIVGMACRFPGVTSAD.....VSPFGVGTNAHVILEAPA 430

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_todent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2265	100.0	3626	Q9F7T9	Q9F7C9 streptomyc
2	2265	100.0	3972	Q9S0R8	Q9S0R8 streptomyc
3	2042	90.2	6239	Q9S0R7	Q9S0R7 streptomyc
4	1968.5	86.9	5532	Q9S0R4	Q9S0R4 streptomyc
5	1744.5	77.0	4881	Q9S0R3	Q9S0R3 streptomyc
6	1733.5	76.5	6048	Q93H87	Q93H87 streptomyc
7	1704	75.2	3170	Q9ALM4	Q9ALM4 saccharopol
8	1700.5	75.1	6145	Q93H84	Q93H84 streptomyc
9	1672	73.8	4928	Q9ALM3	Q9ALM3 saccharopol
10	1668.5	73.1	3524	Q93H86	Q93H86 streptomyc
11	1656.5	73.1	3352	Q93H83	Q93H83 streptomyc
12	1656	73.1	11096	Q914W3	Q914W3 streptomyc
13	1646.5	72.7	4150	Q93H84	Q93H84 streptomyc
14	1645.5	72.6	6797	Q9X993	Q9X993 streptomyc
15	1641	72.5	10917	Q93NM6	Q93NM6 streptomyc
16	1640	72.4	5588	Q9ALM2	Q9ALM2 saccharopol

17	1638	72.3	2723	2	Q30479	Q30479 streptomyc
18	1636	72.2	5060	2	O52545	O52545 amycolacops
19	1636	72.2	5069	2	O52789	O52789 amycolacops
20	1634.5	72.2	1835	2	Q93H85	Q93H85 streptomyc
21	1632	72.1	6146	2	Q93H85	Q93H85 streptomyc
22	1631.5	72.0	4472	2	Q93H84	Q93H84 streptomyc
23	1628.5	71.9	3739	2	Q92G14	Q92G14 streptomyc
24	1626.5	71.8	3970	2	Q93H18	Q93H18 streptomyc
25	1625.5	71.8	4340	2	Q30764	Q30764 streptomyc
26	1622.5	71.6	2595	2	Q9ALM6	Q9ALM6 saccharopol
27	1622.5	71.6	4613	2	Q92G15	Q92G15 streptomyc
28	1619.5	71.5	2152	2	Q9ALM5	Q9ALM5 saccharopol
29	1614.5	71.3	9477	2	Q914X3	Q914X3 streptomyc
30	1613.5	71.2	3576	16	Q9EX53	Q9EX53 streptomyc
31	1613.5	71.2	4498	2	Q93H85	Q93H85 streptomyc
32	1611	71.1	4735	2	O54666	O54666 amycolacops
33	1609	71.0	3816	2	Q9KIV3	Q9KIV3 streptomyc
34	1608.5	71.0	9510	2	Q93H89	Q93H89 streptomyc
35	1606	70.9	1265	2	Q9F847	Q9F847 amycolacops
36	1605	70.9	7525	2	Q9KIE0	Q9KIE0 streptomyc
37	1604.5	70.8	3939	2	Q93H83	Q93H83 streptomyc
38	1601.5	70.7	9507	2	Q9EW41	Q9EW41 streptomyc
39	1597	70.5	3729	2	Q93H86	Q93H86 streptomyc
40	1594.5	70.4	1839	2	Q30765	Q30765 streptomyc
41	1594.5	70.4	2066	2	Q914X1	Q914X1 streptomyc
42	1593.5	70.4	5435	2	Q914X2	Q914X2 streptomyc
43	1591.5	70.3	3413	2	O54593	O54593 amycolacops
44	1590	70.2	7576	2	Q9ZG44	Q9ZG44 streptomyc
45	1589.5	70.2	3654	2	Q30766	Q30766 streptomyc

ALIGNMENTS

RESULT 1

Q9F7T9 PRELIMINARY; PRT; 3626 AA.

AC Q9F7T9; 01-MAR-2001 (TREMUREL_16, Created)
 DT 01-MAR-2001 (TREMUREL_16, Last sequence update)
 DT 01-JUN-2002 (TREMUREL_21, Last annotation update)
 DE Avermectin polypeptide synthase (fragment).
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycesaceae; Streptomyces.
 CX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC31271;
 RA Hong Y.-S., Lee J.J.
 RT "Targeted Gene Disruption of the avermectin O-methyltransferase gene
 and polypeptide synthase gene from Streptomyces avermitilis.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF275943; AAC09812.1; -
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003860; Pantene attach.
 DR Pfam; PF00658; Acyl_transfer; 3.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF00501; ketoacyl-synt C; 2.
 DR PROSITE; PSS0075; ACP DOMAIN; 2.
 DR PROSITE; PSS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE; PSS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 FT NON_TER 3626 3626
 SQ SEQUENCE 3626 AA; 380557 MW; 6272F5F088C1A8D0 CRC64;

Query Match 100.0%; Score 2265; DB 2; Length 3626;
 Best Local Similarity 100.0%; Pred. No. 3.2e-119;
 Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPIAIVGMACRFPGVTSADDFWDLISSEDAIGGFTTGMDLTYDDPDHPGTCYT 60
 |||

Db 2053 EPIAIVGACRPGCVTSADDFWDLISSEODAI GGFPTDRGMDLDTLYDDPDHPTCYT 2112
QY 61 RRGFLVDAGHPDAEFFGSI SPREALAMPPOORLLETAMETIEHAGINPHTLHGTPTGVF 120
Db 2113 RRGFLVDAGHPDAEFFGSI SPREALAMPPOORLLETAMETIEHAGINPHTLHGTPTGVF 2172
QY 121 TGTNGODYALRVHNAAGSTDPGALVTAGSVISGRISYTFGFGPAVSVDTACSSSLVAL 180
Db 2173 TGTNGODYALRVHNAAGSTDPGALVTAGSVISGRISYTFGFGPAVSVDTACSSSLVAL 2232
QY 181 HLAOCALRAECSCMALAGVTVMSPPAFVFEFSRGRGLAADGCKAFSAADGTGWGEGV 240
Db 2233 HLAOCALRAECSCMALAGVTVMSPPAFVFEFSRGRGLAADGCKAFSAADGTGWGEGV 2292
QY 241 GMLIVERLSDAHRNGHRLVAVRGSANVQDASNGLTAPNGPSOQVITROLANAGLSAG 300
Db 2293 GMLIVERLSDAHRNGHRLVAVRGSANVQDASNGLTAPNGPSOQVITROLANAGLSAG 2352
QY 301 DVDAVEAHGTGTTLGDPTEAQAALLATYGGDRAGEGPLMLGSVSNVGHTOAAAGVAGVIK 360
Db 2353 DVDAVEAHGTGTTLGDPTEAQAALLATYGGDRAGEGPLMLGSVSNVGHTOAAAGVAGVIK 2412
QY 361 MWMLRRLGLPRTLHVDESPHYDWSAGAVOLLTETVPMGEGRLRRAVSSFGVSGTN 420
Db 2413 MWMLRRLGLPRTLHVDESPHYDWSAGAVOLLTETVPMGEGRLRRAVSSFGVSGTN 2472
QY 421 AHVILEEAPA 430
Db 2473 AHVILEEAPA 2482

RESULT 2

Q9SOR8 PRELIMINARY; PRT; 3972 AA.

AC Q9SOR8
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 1.
GN AYBA1.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99380548; PubMed=10449723;
RA Ikeda H., Nonomura T., Usami M., Ohta T., Omura S.;
RT "Organization of the biosynthetic gene cluster for the polyketide
RT anheimitic macrolide avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL; AB032367; BA84474.1.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantane attach.
DR InterPro; IPR000834; Zn_carbOdept.
DR Pfam; PF00698; Acyl_transf. 3.
DR Pfam; PF00109; ketoacyl-synt. 2.
DR Pfam; PF02801; ketoacyl-synt. C; 2.
DR Pfam; PF00550; pp-binding; 3.
DR PROSITE; PS50075; ACP DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_2.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 3972 AA; 416852 MW; 2A293659B032B1C3 CRC64;

Query Match 100.0%; Score 2265; DB 2; Length 3972;
Best Local Similarity 100.0%; Pred. No. 3.6e-119;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 EPIAIVGACRPGCVTSADDFWDLISSEODAI GGFPTDRGMDLDTLYDDPDHPTCYT 60
|||||

Db 2086 EPIAIVGACRPGCVTSADDFWDLISSEODAI GGFPTDRGMDLDTLYDDPDHPTCYT 2145
QY 61 RRGFLVDAGHPDAEFFGSI SPREALAMPPOORLLETAMETIEHAGINPHTLHGTPTGVF 120
Db 2146 RRGFLVDAGHPDAEFFGSI SPREALAMPPOORLLETAMETIEHAGINPHTLHGTPTGVF 2205
QY 121 TGTNGODYALRVHNAAGSTDPGALVTAGSVISGRISYTFGFGPAVSVDTACSSSLVAL 180
Db 2206 TGTNGODYALRVHNAAGSTDPGALVTAGSVISGRISYTFGFGPAVSVDTACSSSLVAL 2265
QY 181 HLAOCALRAECSCMALAGVTVMSPPAFVFEFSRGRGLAADGCKAFSAADGTGWGEGV 240
Db 2266 HLAOCALRAECSCMALAGVTVMSPPAFVFEFSRGRGLAADGCKAFSAADGTGWGEGV 2225
QY 241 GMLIVERLSDAHRNGHRLVAVRGSANVQDASNGLTAPNGPSOQVITROLANAGLSAG 300
Db 2226 GMLIVERLSDAHRNGHRLVAVRGSANVQDASNGLTAPNGPSOQVITROLANAGLSAG 2385
QY 301 DVDAVEAHGTGTTLGDPTEAQAALLATYGGDRAGEGPLMLGSVSNVGHTOAAAGVAGVIK 360
Db 2386 DVDAVEAHGTGTTLGDPTEAQAALLATYGGDRAGEGPLMLGSVSNVGHTOAAAGVAGVIK 2445
QY 361 MWMLRRLGLPRTLHVDESPHYDWSAGAVOLLTETVPMGEGRLRRAVSSFGVSGTN 420
Db 2446 MWMLRRLGLPRTLHVDESPHYDWSAGAVOLLTETVPMGEGRLRRAVSSFGVSGTN 2505
QY 421 AHVILEEAPA 430
Db 2506 AHVILEEAPA 2515

RESULT 3

Q9SOR7 PRELIMINARY; PRT; 6239 AA.

AC Q9SOR7
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 2.
GN AYBA2.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99380548; PubMed=10449723;
RA Ikeda H., Nonomura T., Usami M., Ohta T., Omura S.;
RT "Organization of the biosynthetic gene cluster for the polyketide
RT anheimitic macrolide avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL; AB032367; BA84475.1.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantane attach.
DR InterPro; IPR000169; SHprot acsite.
DR Pfam; PF00698; Acyl_transf. 4.
DR Pfam; PF00109; ketoacyl-synt. 4.
DR Pfam; PF02801; ketoacyl-synt. C; 4.
DR Pfam; PF00550; pp-binding; 4.
DR PROSITE; PS50075; ACP DOMAIN; 4.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 4.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_4.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 3.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 6239 AA; 666283 MW; A706280940B502FA CRC64;

Query Match 90.2%; Score 2042; DB 2; Length 6239;
Best Local Similarity 89.5%; Pred. No. 2.4e-106;
Matches 385; Conservative 14; Mismatches 31; Indels 0; Gaps 0;


```

Qy 1 EPIAIVGMACRFPGCVTSADDFMDLISSEODAI GGFPTDRGMDLDTLYDDPDHPTCYT 60
Db 2680 EPIAIVGMACRFPGCVTSADDFMDLISSEODAI GGFPTDRGMDLDTLYDDPDHPTCYT 2739
Qy 61 RNGCFLYDAGHFDPAEPFGISPREALAMDPOORLLLETAMETIEHAGINPHTLHGTPTGVF 120
Db 2740 RNGCFLYDAGHFDPAEPFGISPREALAMDPOORLLLETAMETIEHAGINPHTLHGTPTGVF 2799
Qy 121 TGTNGODYALRVH-NAGOSTDGFALTGTAAGVSI SGRISYTFEGEPGPAVSVDTCSSSLVA 180
Db 2800 TGTNGODYALRVH-NAGOSTDGFALTGTAAGVSI SGRISYTFEGEPGPAVSVDTCSSSLVA 2859
Qy 181 HLAACQALRAGECSMALAGVTVMS SPGAFVEFSRQGLADGHCASFSAADGTGMEGV 240
Db 2860 HLAACQALRAGECSMALAGVTVMS SPGAFVEFSRQGLADGHCASFSAADGTGMEGV 2919
Qy 241 GMLLVERLSDAHRNGHRLVAVRGSANODGASNGITAPNGSQORVITQALANAGLSA 300
Db 2920 GMLLVERLSDAHRNGHRLVAVRGSANODGASNGITAPNGSQORVITQALANAGLSA 2979
Qy 301 DVDAVEAHGCTGTTLDGPPIEAQALLATYGGDRAGEGFLMGSKVSNVGHTOAAGVAVGI 360
Db 2980 DVDAVEAHGCTGTTLDGPPIEAQALLATYGGDRAGEGFLMGSKVSNVGHTOAAGVAVGI 3039
Qy 361 MVMALRHGILPRTLHVDEBSPHVDMSAGAVQLLTETVPMWPGEGRLRAGVSSFGVSGTN 420
Db 3040 MVMALRHGILPRTLHVDEBSPHVDMSAGAVQLLTETVPMWPGEGRLRAGVSSFGVSGTN 3099
Qy 421 AHVILEEAPA 430
Db 3100 AHVILEEAPA 3109

```

RESULT 4
Q9S0R4 PRELIMINARY; PRT; 5532 AA.

```

AC Q9S0R4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 3.
GN AVEA3.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9380548; PubMed=10449723;
RA Ikeda H., Nonomiyama T., Usami M., Ohta T., Omura S.;
RT "Organization of the biosynthetic gene cluster for the polyketide
RT anthelmintic macroide avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL; AB032367; BAA84478.1; -.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantene-attach.
DR InterPro; IPR000834; Zn carbodepe.
DR Pfam; PF00698; Acyl_transf; 3.
DR Pfam; PF00109; ketoacyl-synt; 3.
DR Pfam; PF02801; ketoacyl-synt_C; 3.
DR Pfam; PF00550; pp-binding; 3.
DR PROSITE; PS50075; ACP DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOMN_1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
DR Phosphopantetheine; Transferase.
SQ SEQUENCE 5532 AA; 575193 MW; EC7B2069DD4A4057 CRC64;

```

Query Match 86.9%; Score 1968.5; DB 2; Length 5532;
Best Local Similarity 87.2%; Pred. No. 2.9e-102;
Matches 376; Conservative 23; Mismatches 31; Indels 1; Gaps 1;

```

Qy 1 EPIAIVGMACRFPGCVTSADDFMDLISSEODAI GGFPTDRGMDLDTLYDDPDHPTCYT 60
Db 1864 EPIAIVGMACRFPGCVTSADDFMDLISSEODAI GGFPTDRGMDLDTLYDDPDHPTCYT 1923
Qy 61 RNGCFLYDAGHFDPAEPFGISPREALAMDPOORLLLETAMETIEHAGINPHTLHGTPTGVF 120
Db 1924 RNSFLYDAGHFDPAEPFGISPREALAMDPOORLLLETAMETIEHAGINPHTLHGTPTGVF 1983
Qy 121 TGTNGODYALRVH-NAGOSTDGFALTGTAAGVSI SGRISYTFEGEPGPAVSVDTCSSSLVA 179
Db 1984 AGILPQAYGRLEHMAADTEGYVLTGSSVASGRISYTFEGEPGPAVSVDTCSSSLVA 2043
Qy 180 HLAACQALRAGECSMALAGVTVMS SPGAFVEFSRQGLADGHCASFSAADGTGMEGV 239
Db 2044 HLAACQALRAGECSMALAGVTVMS SPGAFVEFSRQGLADGHCASFSAADGTGMEGV 2103
Qy 240 GMLLVERLSDAHRNGHRLVAVRGSANODGASNGITAPNGSQORVITQALANAGLSA 299
Db 2104 GMLLVERLSDAHRNGHRLVAVRGSANODGASNGITAPNGSQORVITQALANAGLSA 2163
Qy 300 DVDAVEAHGCTGTTLDGPPIEAQALLATYGGDRAGEGFLMGSKVSNVGHTOAAGVAVGI 359
Db 2164 DVDAVEAHGCTGTTLDGPPIEAQALLATYGGDRAGEGFLMGSKVSNVGHTOAAGVAVGI 2223
Qy 360 MVMALRHGILPRTLHVDEBSPHVDMSAGAVQLLTETVPMWPGEGRLRAGVSSFGVSGTN 419
Db 2224 MVMALRHGILPRTLHVDEBSPHVDMSAGAVQLLTETVPMWPGEGRLRAGVSSFGVSGTN 2283
Qy 420 NAHVILEEAPA 430
Db 2284 NAHVILEEAPA 2294

```

RESULT 5
Q9S0R3 PRELIMINARY; PRT; 4881 AA.

```

AC Q9S0R3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 4.
GN AVEA4.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9380548; PubMed=10449723;
RA Ikeda H., Nonomiyama T., Usami M., Ohta T., Omura S.;
RT "Organization of the biosynthetic gene cluster for the polyketide
RT anthelmintic macroide avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL; AB032367; BAA84479.1; -.
DR HSP; P25715; IMLA.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantene-attach.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00698; Acyl_transf; 3.
DR Pfam; PF00109; ketoacyl-synt; 3.
DR Pfam; PF02801; ketoacyl-synt_C; 3.
DR Pfam; PF00550; pp-binding; 3.
DR PROSITE; PS50075; Thioesterase; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
DR Phosphopantetheine; Transferase.
SQ SEQUENCE 4881 AA; 510303 MW; F6568C1F01ACD37A CRC64;

```

Query Match 77.0%; Score 1744.5; DB 2; Length 4881;
Best Local Similarity 76.7%; Pred. No. 1.1e-89;
Matches 339; Conservative 40; Mismatches 59; Indels 1; Gaps 1;

```

QY 1 EPIAIVGMACRPFPGVTSADDFWDLISSEODALGFPPTDRGMDLTYDPPDPHGCYCT 60
2636 EPIAIVGMACRPFPGVTSADDFWDLISSEODALGFPPTDRGMDLTYDPPDPHGCYCT 2635
Db 61 RRGFLYDAGHPDAEFGISPREALAMPQORLLLETAMETIEHAGINPTLHGTPTGVF 120
2696 RRGFLYDAGHPDAEFGISPREALAMPQORLLLETAMETIEHAGINPTLHGTPTGVF 2755
QY 121 TGTNGODYALRVNNAAGSTDFALGTAGSVISGRISYTFEGFEPGPAVSDTACSSSLVAL 180
2756 AGVSGQDYALRVNNAAGSTDFALGTAGSVISGRISYTFEGFEPGPAVSDTACSSSLVAL 2815
Db 181 HLAQALRAGECSMALAGVTVMSPGAFAVEFSRQGLADGCKAFSAADGTGMEGV 240
2816 HLAQALRAGECSMALAGVTVMSPGAFAVEFSRQGLADGCKAFSAADGTGMEGV 2875
QY 241 GMLIVERLSDAHRNGHRYLAVRGSANVODGASNGLTAPNGPQORVIRQALANAGLSAG 300
2876 GMLIVERLSDAHRNGHRYLAVRGSANVODGASNGLTAPNGPQORVIRQALANAGLSAG 2935
QY 301 DVDAVEAHGTGTTGDPTEAQLLATYGGDRAGEPLMLGSKVSNVGHGTOAAAGVAVIK 360
2936 DVDAVEAHGTGTTGDPTEAQLLATYGGDRAGEPLMLGSKVSNVGHGTOAAAGVAVIK 2995
QY 361 MVMALRHGLLPTLHVDPEPSPHVDMSAGAVOLLTEVPWPGEGRLRRAGVSSFGVSGTN 420
2996 MVMALRHGLLPTLHVDPEPSPHVDMSAGAVOLLTEVPWPGEGRLRRAGVSSFGVSGTN 3054
Db 421 AHVILEEAP 429
3055 AHVILEEAP 3063
QY 421 AHVILEEAP 429
3055 AHVILEEAP 3063

```

RESULT 6

Q93H87 PRELIMINARY; PRT; 6048 AA.

```

ID Q93H87 PRELIMINARY; PRT; 6048 AA.
AC Q93H87;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Modular polyketide synthase.
GN PTEA1.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyicinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinohe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: Deducing the ability of producing secondary
RT metabolites.";
RL EMBL: AB070949; BAB69303.1; -
DR EMBL: IPR001227; Ac transferase.
DR InterPro: IPR004410; FAD.
DR InterPro: IPR000794; Ketocacyl-synt.
DR InterPro: IPR003880; Pantne attach.
DR Pfam: PF00698; Acyl transf. 4.
DR Pfam: PF00109; ketocacyl-synt. 4.
DR Pfam: PF02801; ketocacyl-synt. 4.
DR Pfam: PF00550; pp-binding. 4.
DR TIGRFAMs: TIGR00128; fadD. 4.
DR PROSITE: PS50075; ACP DOMAIN. 4.
DR PROSITE: PS00606; B KETOACYL SYNTHASE; UNKNOWN 3.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN 4.
KW Phosphopantetheine.
SQ SEQUENCE 6048 AA; 629420 MW; F8673CE2D2694705 CRC64;

```

Query Match 76.5%; Score 1733.5; DB 2; Length 6048;
 Best Local Similarity 76.3%; Pred. No. 5.8e-89;
 Matches 328; Conservative 42; Mismatches 59; Indels 1; Gaps 1;

```

QY 1 EPIAIVGMACRPFPGVTSADDFWDLISSEODALGFPPTDRGMDLTYDPPDPHGCYCT 60
2552 EPIAIVGMACRPFPGVTSADDFWDLISSEODALGFPPTDRGMDLTYDPPDPHGCYCT 2611
Db 61 RRGFLYDAGHPDAEFGISPREALAMPQORLLLETAMETIEHAGINPTLHGTPTGVF 120
2612 RRGFLYDAGHPDAEFGISPREALAMPQORLLLETAMETIEHAGINPTLHGTPTGVF 2671
QY 121 TGTNGODYALRVNNAAGSTDFALGTAGSVISGRISYTFEGFEPGPAVSDTACSSSLVAL 180
2672 TGTNGODYALRVNNAAGSTDFALGTAGSVISGRISYTFEGFEPGPAVSDTACSSSLVAL 2731
Db 181 HLAQALRAGECSMALAGVTVMSPGAFAVEFSRQGLADGCKAFSAADGTGMEGV 240
2732 HLAQALRAGECSMALAGVTVMSPGAFAVEFSRQGLADGCKAFSAADGTGMEGV 2791
QY 241 GMLIVERLSDAHRNGHRYLAVRGSANVODGASNGLTAPNGPQORVIRQALANAGLSAG 300
2792 GMLIVERLSDAHRNGHRYLAVRGSANVODGASNGLTAPNGPQORVIRQALANAGLSAG 2851
QY 301 DVDAVEAHGTGTTGDPTEAQLLATYGGDRAGEPLMLGSKVSNVGHGTOAAAGVAVIK 360
2852 DVDAVEAHGTGTTGDPTEAQLLATYGGDRAGEPLMLGSKVSNVGHGTOAAAGVAVIK 2911
QY 361 MVMALRHGLLPTLHVDPEPSPHVDMSAGAVOLLTEVPWPGEGRLRRAGVSSFGVSGTN 420
2912 MVMALRHGLLPTLHVDPEPSPHVDMSAGAVOLLTEVPWPGEGRLRRAGVSSFGVSGTN 2970
Db 421 AHVILEEAP 430
2971 AHVILEEAP 2980
QY 421 AHVILEEAP 430
2971 AHVILEEAP 2980

```

RESULT 7

Q9ALM4 PRELIMINARY; PRT; 3170 AA.

```

ID Q9ALM4 PRELIMINARY; PRT; 3170 AA.
AC Q9ALM4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polyketide synthase extender modules 3-4.
GN SPNC.
OS Saccharopolyspora spinosa.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;
OC Saccharopolyspora.
OX NCBI_TaxID=60894;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21257765; PubMed=11358695;
RA Waldron C., Matsushima P., Rostek P.R. Jr., Broughton M.C.,
RA Turner J., Madduri K., Crawford K.P., Merlo D.J., Baltz R.H.;
RT "Cloning and analysis of the spinosad biosynthetic cluster of
RT Saccharopolyspora spinosa.";
RL Chem. Biol. 8:487-499(2001).
DR EMBL: AY007564; AAG23266.1; -.
DR HSSP: P25715; IMLA.
DR InterPro: IPR001227; Ac transferase.
DR InterPro: IPR002198; ADH short.
DR InterPro: IPR000794; Ketocacyl-synt.
DR InterPro: IPR003880; Pantne attach.
DR Pfam: PF00698; Acyl transf. 2.
DR Pfam: PF00109; ketocacyl-synt. 2.
DR Pfam: PF02801; ketocacyl-synt. 2.
DR Pfam: PF00550; pp-binding. 2.
DR PROSITE: PS50075; ACP DOMAIN. 2.
DR PROSITE: PS00606; B KETOACYL SYNTHASE; 2.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 2.

```

KW Phosphopantetheine; Transferase.
SQ SEQUENCE 3170 AA; 334873 MW; 205D519FDEA3249C CRC64;

Query Match 75.2%; Score 1704; DB 2; Length 3170;
Best Local Similarity 75.3%; Pred. No. 1,2e-87;

Matches 323; Conservative 38; Mismatches 66; Indels 2; Gaps 1;

QY 1 EPIAIVGMACRFPGVTSADDFMDLISSEDAIGFPTDRGMDLDTLYDDPDPHGTCYT 60
DB 34 EPIAIVGMACRFPGVTSADDFMDLISSEDAIGFPTDRGMDLDTLYDDPDPHGTCYT 93
QY 61 RINGFLYDAGHFAEFPGISPREALAMPQORLLLETAMETIHAGINPHTLTGTPGVF 120
DB 94 RHGGFLHDAGFDFDIPDISPREALAMPQORLLLETAMETIHAGINPHTLTGTPGVF 153
QY 121 TGTNGODYALRVHNAAGSTGDFALTGAGSVISGRISTGFEFGPAVSVDPAACSSSLVAL 180
DB 154 AGLMYEGYDTGAMHAGSGVSEGLTGAGSVASGRVAVAFEGECPAVTVDPAACSSSLVAL 213
QY 181 HLAQALRAGECSMALAGVTVMSPGAFFEFGRGLAADGHCASFSAADGTGMEGV 240
DB 214 HLAQALRAGECSMALAGVTVMSPGAFFEFGRGLAADGHCASFSAADGTGMEGV 273
QY 241 GMLIVERLSDAHNRHGRVLAIVRGSAVNOGASNGLTAPNGPQOQRYTQALANAGLSAG 300
DB 274 GLVLERLSDAHNRHGRVLAIVRGSAVNOGASNGLTAPNGPQOQRYTQALANAGLSAG 333
QY 301 DVDAVEAHGCTTLDGPIEAOALATYGGORABEGPLMGSVKSNVNGHTQAAAGVAVIK 360
DB 334 DVDAVEAHGCTTLDGPIEAOALATYGGORABEGPLMGSVKSNVNGHTQAAAGVAVIK 393
QY 361 MVNALRHGLLPRTLHVDSPPHVMSAGAVOLLTETVPMPGEGRLRRAGVSFGVSGTN 420
DB 394 MVNAMRHGGLPRTLHVDSPPHVMSAGAVOLLTETVPMPGEGRLRRAGVSFGVSGTN 451
QY 421 AHVILEAP 429
DB 452 AHVILEHDP 460

RESULT 8
Q93H84 PRELIMINARY; PRT; 6145 AA.
AC Q93H84;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Modular polyketide synthase.
GN PTEA4.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Iehikawa J., Hanamoto A., Takahashi C.,
RA Shinozaki M., Takahashi Y., Horikawa H., Nakazawa H., Osorio T.,
RA Kikuchi H., Shida T., Sakai Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: Deducing the ability of producing secondary
RT metabolites";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; AB070949; BAB69306.1; -
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR004410; FAD.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR000566; Lipocin_cyFABP.
DR InterPro; IPR003880; Pantetheine.
DR Pfam; PF00698; Acyl_transf_4.
DR Pfam; PF00109; Ketoacyl-synt; 4.
DR Pfam; PF02801; Ketoacyl-synt_C; 4.
DR Pfam; PF00550; pp-binding; 4.

DR TIGRFAM; TIGR00128; fad; 4.
DR PROSITE; PS50075; ACP DOMAIN; 4.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_4.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_4.
KW Phosphopantetheine.
SQ SEQUENCE 6145 AA; 642013 MW; 6A75B1E7A7DE95 CRC64;

Query Match 75.1%; Score 1700.5; DB 2; Length 6145;
Best Local Similarity 75.5%; Pred. No. 4,3e-87;
Matches 324; Conservative 39; Mismatches 65; Indels 1; Gaps 1;

QY 1 EPIAIVGMACRFPGVTSADDFMDLISSEDAIGFPTDRGMDLDTLYDDPDPHGTCYT 60
DB 4576 EPIAIVGMACRFPGVTSADDFMDLISSEDAIGFPTDRGMDLDTLYDDPDPHGTCYT 4635
QY 61 RINGFLYDAGHFAEFPGISPREALAMPQORLLLETAMETIHAGINPHTLTGTPGVF 120
DB 4636 RINGFLYDAGHFAEFPGISPREALAMPQORLLLETAMETIHAGINPHTLTGTPGVF 4695
QY 121 TGTNGODYALRVHNAAGSTGDFALTGAGSVISGRISTGFEFGPAVSVDPAACSSSLVAL 180
DB 4696 VGSQYDYLARARNAIDDSGFFGTGNSASIMSGRIATTTGLGPAVTVDTACSSSLVAL 4755
QY 181 HLAQALRAGECSMALAGVTVMSPGAFFEFGRGLAADGHCASFSAADGTGMEGV 240
DB 4756 HMAIQALRAGECSMALAGVTVMSPGAFFEFGRGLAADGHCASFSAADGTGMEGV 4815
QY 241 GMLIVERLSDAHNRHGRVLAIVRGSAVNOGASNGLTAPNGPQOQRYTQALANAGLSAG 300
DB 4816 GMLIVERLSDAHNRHGRVLAIVRGSAVNOGASNGLTAPNGPQOQRYTQALANAGLSAG 4875
QY 301 DVDAVEAHGCTTLDGPIEAOALATYGGORABEGPLMGSVKSNVNGHTQAAAGVAVIK 360
DB 4876 DVDAVEAHGCTTLDGPIEAOALATYGGORABEGPLMGSVKSNVNGHTQAAAGVAVIK 4935
QY 361 MVNALRHGLLPRTLHVDSPPHVMSAGAVOLLTETVPMPGEGRLRRAGVSFGVSGTN 420
DB 4936 MVNAMRHGGLPRTLHVDSPPHVMSAGAVOLLTETVPMPGEGRLRRAGVSFGVSGTN 4994
QY 421 AHVILEAP 429
DB 4995 AHVILEOAP 5003

RESULT 9
Q9ALM3 PRELIMINARY; PRT; 4928 AA.
AC Q9ALM3;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polyketide synthase extender modules 5-7.
GN SPND.
OS Saccharopolyspora spinosa.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Pseudonocardiales; Pseudonocardaceae;
OC Saccharopolyspora.
OX NCBI_TaxID=60894;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21257765; PubMed=1158695;
RA Waldron C., Matsushima P., Rosebeck P.R. Jr., Broughton M.C.,
RA Turner J., Madduri K., Crawford K.P., Merlo D.J., Baltz R.H.;
RT "Cloning and analysis of the spinosad biosynthetic cluster of
RT Saccharopolyspora spinosa";
RL Chem. Biol. 8:487-499(2001).
DR EMBL; AY007564; AAG23263.1; -
DR HSP; P25715; IMIA.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR004410; FAD.
DR InterPro; IPR000794; Ketoacyl-synt.

DR InterPro; IPR003880; Pantne attach.
 DR Pfam; PF00698; Acyl_transf_3.
 DR Pfam; PF00106; adh_thort; 1.
 DR Pfam; PF00109; ketoacyl-synt; 3.
 DR Pfam; PF02801; ketoacyl-synt_C; 3.
 DR Pfam; PF00550; pp-binding; 3.
 DR TIGRfam; TIGR00128; fabD; 3.
 DR PROSITE; PSS0075; ACP DOMAIN; 3.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
 DR Phosphopantetheine; Transferase.
 KM SEQUENCE 4928 AA; 517247 MW; C22C6361888F865A CRC64;
 Query Match 73.8%; Score 1672; DB 2; Length 4928;
 Best Local Similarity 75.2%; Pred. No. 1.3e-85;
 Matches 322; Conservative 34; Mismatches 70; Indels 2; Gaps 2;
 QY 3 IAIIVGACRPPGVTSADDFWDLISSEBDALGFPPTDRGMDLDTLYDPDPHPGTCYTRN 62
 DB 1787 VAIIVGRCRPPGVTSPEELRLVAGVDVAGFPDDRGMDLALYDPDPRLGTSTVCE 1846
 QY 63 GGFLYDAGHPAEFFGISPREALAMPQORLLLETAMETTHAGINPHTLHGTPTGVTG 122
 DB 1847 GGFRLDAEFPDAMFGISPREALAMPQORLLLETAMETTERAGIDPFLHGSRTGVFAG 1906
 QY 123 TNGGDYALR-VHNAAGSTDFALGTAGTSVSGRISYTFEGEPGAVSVDTACSSSLVALH 181
 DB 1907 LMYHDYGRFTRTPRPGEGHGLGTGNAGSVLSGRVASFEGEPGAVVDTACSSSLVALH 1966
 QY 182 LACQALRAGECSMALAGGVTVSSPGAFVFEFSRQGLAADGCKAFSAADGTGMEGVG 241
 DB 1967 LAQALRAGECEPFLAGGVTVSSPTTFEFSRQGLAADGCKAFSAADGTGMEGVG 2026
 QY 242 MLIVERSDARNRHRVLAIVRGSANVDGASNGLTAPNGSQORVIRQALANGLSAGD 301
 DB 2027 LVLLERSDARNRHRVLAIVRGSANVDGASNGLTAPNGSQORVIRQALANGLSAGD 2086
 QY 302 VDAVEAHGTTGTLGPIEAQALLATYGGDRAGEGFLMGVSKSVNGHTQAAAGVATKM 361
 DB 2087 VDAVEAHGTTGTLGPIEAQALLATYGGDRAGEGFLMGVSKSVNGHTQAAAGVATKM 2146
 QY 362 VMALRHGLPRTLHVDEPSPHVDWSAGAVQLLTETVPWPGEGRLRRAGVSPFVSGTNA 421
 DB 2147 VMARQGLPRTLHVDEPSAGVDWSAGAVQLLTETVPWPGEGRLRRAGVSPFVSGTNA 2205
 QY 422 HVILIEEAP 429
 DB 2206 HLIEQEP 2213
 RESULT 10
 Q93H86 PRELIMINARY; PRT; 3524 AA.
 AC Q93H86;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Modular polyketide synthase.
 GN PTEA2.
 OS Streptomyces avermiltilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 avermiltilis: Deducing the ability of producing secondary
 metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).

DR EMBL; AB070949; BAB69304.1; -
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR001993; Mitoch_carrier.
 DR InterPro; IPR003880; Pantne attach.
 DR Pfam; PF00698; Acyl_transf_3.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF02801; ketoacyl-synt_C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PSS0075; ACP DOMAIN; 2.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_2.
 KM Phosphopantetheine.
 SO SEQUENCE 3524 AA; 366619 MW; 64A65759C461EBCA CRC64;
 Query Match 73.7%; Score 1668.5; DB 2; Length 3524;
 Best Local Similarity 74.1%; Pred. No. 1.4e-85;
 Matches 318; Conservative 40; Mismatches 70; Indels 1; Gaps 1;
 QY 1 EPIAIVGACRPPGVTSADDFWDLISSEBDALGFPPTDRGMDLDTLYDPDPHPGTCYTRN 60
 DB 1730 EPIAIVGACRPPGVTSADDFWDLISSEBDALGFPPTDRGMDLDTLYDPDPHPGTCYTRN 1789
 QY 61 RNCGLYDAGHPAEFFGISPREALAMPQORLLLETAMETTHAGINPHTLHGTPTGVTG 120
 DB 1790 RNCGLYDAGHPAEFFGISPREALAMPQORLLLETAMETTERAGIDPFLHGSRTGVF 1849
 QY 121 TNGGDYALR-VHNAAGSTDFALGTAGTSVSGRISYTFEGEPGAVSVDTACSSSLVALH 180
 DB 1850 TNGGDYALR-VHNAAGSTDFALGTAGTSVSGRISYTFEGEPGAVSVDTACSSSLVALH 1909
 QY 181 HLAQALRAGECSMALAGGVTVSSPGAFVFEFSRQGLAADGCKAFSAADGTGMEGVG 240
 DB 1910 HLAQALRAGECSMALAGGVTVSSPTTFEFSRQGLAADGCKAFSAADGTGMEGVG 2029
 QY 241 GMLIVERSDARNRHRVLAIVRGSANVDGASNGLTAPNGSQORVIRQALANGLSAGD 300
 DB 1970 GMLIVERSDARNRHRVLAIVRGSANVDGASNGLTAPNGSQORVIRQALANGLSAGD 2029
 QY 301 DVDAVEAHGTTGTLGPIEAQALLATYGGDRAGEGFLMGVSKSVNGHTQAAAGVATKM 360
 DB 2030 DVDAVEAHGTTGTLGPIEAQALLATYGGDRAGEGFLMGVSKSVNGHTQAAAGVATKM 2089
 QY 361 MMALRHGLPRTLHVDEPSPHVDWSAGAVQLLTETVPWPGEGRLRRAGVSPFVSGTNA 420
 DB 2090 MMAMRRGLPRTLHVDEPSPHVDWSAGAVQLLTETVPWPGEGRLRRAGVSPFVSGTNA 2148
 QY 421 AHVILIEEAP 429
 DB 2149 AHVILIEEAP 2157
 RESULT 11
 Q93H83 PRELIMINARY; PRT; 3352 AA.
 AC Q93H83;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Modular polyketide synthase.
 GN PTEA5.
 OS Streptomyces avermiltilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces

RT avertmilitis: Deducing the ability of producing secondary
 RT metadollies.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL; AB070949; BAB69307.1; -
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002198; Adh_short.
 DR InterPro; IPR002362; Antenna_beta.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantene_attach.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00698; Acyl_transf. 2.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF02801; ketoacyl-synt C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00975; Thioesterase; 1.
 DR PROSITE; PSS50075; ACP_DOMAIN; 2.
 DR PROSITE; PS00969; ANTENNA_COMP_BETA; UNKNOWN_1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_2.
 KM Phosphopantetheine.
 SQ SEQUENCE 3352 AA; 353395 MW; 66A5865BFD2027B7 CRC64;
 Query Match 73.1%; Score 1656.5; DB 2; Length 3352;
 Best Local Similarity 73.4%; Pred. No. 6.2e-85;
 Matches 315; Conservative 42; Mismatches 71; Indels 1; Gaps 1;
 QY 1 EPIAIVGMACRFPGVTSADDFPMDLISSEDAIGSPFTRDGMWLDITLDPDPHRCYCT 60
 DB 35 DPLAIVGMACRFPGVTSADDFPMDLISSEDAIGSPFTRDGMWLDITLDPDPHRCYCT 94
 QY 61 RRGGLFLVDAGHPDAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLGTPTGVF 120
 DB 95 RRGGLFLVDAGHPDAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLGTPTGVF 154
 QY 121 TGTNGDYALRVHNAQSTGDFALITGAGSVISGRISYTFGEGPAVSVDTACSSSLVAL 180
 DB 155 VGSQYQDYVNLVGVGSESDGHLCTGNSASVMSGRIVTFGLGPAVTVDTCSSSLVAL 214
 QY 181 HLMCOALRAGECSMALAGVTVMSSPGAFVFSRORGLADGHCASAAADGTGMEGV 240
 DB 215 HMAIQALRAGECSMALAGVTVMSSPGAFVFSRORGLADGHCASAAADGTGMEGV 274
 QY 241 GMLLVERLSDAHNRGHRVLAVRGSAVNODGASNGLTAPNGPSQORYIROLANAGLSAG 300
 DB 275 GMLLVERLSDAHNRGHRVLAVRGSAVNODGASNGLTAPNGPSQORYIROLANAGLSAG 334
 QY 301 DVDAVEAHGTTGLGDPPIEAQALLATYGGDRAGEGPIMLGSVKSNGHTQAAAGVAVIK 360
 DB 335 QIDVVEAHGTTGLGDPPIEAQALLATYGGDRAGEGPIMLGSVKSNGHTQAAAGVAVIK 394
 QY 361 MWALRHGGLPRTLHVDSPRPHVDSAGAVQLLTETVPWPGEGRLRRAGVSSFGVSGTN 420
 DB 395 MWAMRHGGLPRTLHVDSPRPHVDSAGAVQLLTETVPWPGEGRLRRAGVSSFGVSGTN 453
 QY 421 AHVILEAP 429
 DB 454 AHVILEAP 462

RESULT 12
 ID 09LAW3 PRELIMINARY; PRT; 11096 AA.
 AC 09LAW3;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE NYSC.
 GN NYSC.
 OS Streptomyces nourset.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 NCBI_TaxId=1971;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 11455;
 RX MEDLINE=20334850; Pubmed=10873841;
 RA Brautaset T., Sekurova O.N., Sletta H., Ellingren T.E., Strom A.R.,
 RA Valla S., Zotchev S.B.;
 RT "Biosynthesis of the polyene antifungal antibiotic nystatin in
 RT Streptomyces nourset ATCC 11455: analysis of the gene cluster and
 RT deduction of the biosynthetic pathway.";
 RL Chem. Biol. 7:395-403(2000).
 DR EMBL; AF263912; AF27176.1; -
 DR HSSP; P25715; IMLA.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh_zn_family.
 DR InterPro; IPR004410; Fadh.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantene_attach.
 DR Pfam; PF00698; Acyl_transf. 6.
 DR Pfam; PF00109; adh_zinc; 1.
 DR Pfam; PF00107; ketoacyl-synt; 6.
 DR Pfam; PF00550; pp-binding; 6.
 DR Pfam; PF02801; ketoacyl-synt C; 6.
 DR TIGRfam; TIGR00128; fadh; 6.
 DR PROSITE; PSS50075; ACP_DOMAIN; 6.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 6.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 5.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 11096 AA; 1150415 MW; 776CAEAFCAE55IDD CRC64;
 Query Match 73.1%; Score 1656; DB 2; Length 11096;
 Best Local Similarity 73.6%; Pred. No. 2.9e-84;
 Matches 320; Conservative 40; Mismatches 67; Indels 8; Gaps 3;
 QY 1 EPIAIVGMACRFPGVTSADDFPMDLISSEDAIGSPFTRDGMWLDITLDPDPHRCYCT 60
 DB 3501 DPLAIVGMACRFPGVTSADDFPMDLISSEDAIGSPFTRDGMWLDITLDPDPHRCYCT 3560
 QY 61 RRGGLFLVDAGHPDAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLGTPTGVF 120
 DB 3561 RRGGLFLVDAGHPDAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLGTPTGVF 3620
 QY 121 TGTNGDYALRVHNAQSTGDFALITGAGSVISGRISYTFGEGPAVSVDTACSSSLVAL 180
 DB 3621 AGWVSDYGSIL--GGKFEFGFGQGSAGSVASGRVYALGFGGPAVTVDTCSSSLVAL 3678
 QY 181 HLMCOALRAGECSMALAGVTVMSSPGAFVFSRORGLADGHCASAAADGTGMEGV 240
 DB 3679 HMAIQALRAGECSMALAGVTVMSSPGAFVFSRORGLADGHCASAAADGTGMEGV 3738
 QY 241 GMLLVERLSDAHNRGHRVLAVRGSAVNODGASNGLTAPNGPSQORYIROLANAGLSAG 300
 DB 3739 GILVLEKSDAVNRGHRVLAVRGSAVNODGASNGLTAPNGPSQORYIROLANAGLSAG 3798
 QY 301 DVDAVEAHGTTGLGDPPIEAQALLATYGGDRAGEGPIMLGSVKSNGHTQAAAGVAVIK 360
 DB 3799 DVDAVEAHGTTGLGDPPIEAQALLATYGGDRAGEGPIMLGSVKSNGHTQAAAGVAVIK 3858
 QY 361 MWALRHGGLPRTLHVDSPRPHVDSAGAVQLLTETVPWPGEGRLRRAGVSSFGVSGTN 420
 DB 3859 MWAMRHGGLPRTLHVDSPRPHVDSAGAVQLLTETVPWPGEGRLRRAGVSSFGVSGTN 3917
 QY 421 AHVILEAP 430
 DB 3918 AHVILEAP 3932

RESULT 13
 ID 09KIV4 PRELIMINARY; PRT; 4150 AA.
 AC 09KIV4;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE 8,8a-deoxyoleandolide synthase 1.
 GN OLEAI.
 OS Streptomyces antibiotics.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1890;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20363406; PubMed=10908114;
 RA Shah S., Xue Q., Tang L., Carney J.R., Betlach M., McDaniel R.;
 RT "Cloning, Characterization, and Heterologous Expression of a
 RT Polyketide Synthase and P-450 Oxidase Involved in the Biosynthesis of
 RT the Antibiotic Oleandomycin.";
 RL J. Antibiot. 53:502-506(2000).
 DR EMBL: AF220951; AAF2408.1; -
 DR HSSP; P25715; IMLA.
 DR InterPro; IPR002106; ALCRNA_ligase1.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR004410; FabD.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantone attach.
 DR InterPro; IPR000169; Shpox_acsice.
 DR Pfam; PF00698; Acyl_transf; 3.
 DR Pfam; PF02801; ketoacyl-synt; 3.
 DR Pfam; PF00550; pp-binding; 3.
 DR TIGRfams; TIGR00128; fabd; 3.
 DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 DR PROSITE; PS00075; ACP_DOMAIN; 3.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 3.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
 KW Phosphopantetheine; Transferase.
 SQ SEQUENCE 4150 AA; 435261 MW; 9383296C4C16667D CRC64;

Query Match 72.7%; Score 1646.5; DB 2; Length 4150;
 Best Local Similarity 72.8%; Pred. No. 3e-84;
 Matches 313; Conservative 41; Mismatches 75; Indels 1; Gaps 1;

QY 1 EPPIAVGACRFGPGVTSADDFMDLISSEDAIGGFPFDRGMDLDTLYDPDPHPTGCT 60
 DB EPIAVGACRFGPGVTSADDFMDLISSEDAIGGFPFDRGMDLDTLYDPDPHPTGCT 60
 2546 EPIAVGACRFGPGVTSADDFMDLISSEDAIGGFPFDRGMDLDTLYDPDPHPTGCT 60
 QY 61 RNGEFLYDAGHPDAEFPGISPREALAMPPOORLLLETAMETIEHAGINPHTLHGTPTGVF 120
 DB REGAFLEHAGFPAEFPGISPREALAMPPOORLLLETAMETIEHAGINPHTLHGTPTGVF 120
 2606 REGAFLEHAGFPAEFPGISPREALAMPPOORLLLETAMETIEHAGINPHTLHGTPTGVF 120
 QY 121 TGTNGODYALRVHNAAGSTGDFALTTAGSVISGRISYTFEGEPGPAVSVDTCSSSLVAL 180
 DB TGTNGODYALRVHNAAGSTGDFALTTAGSVISGRISYTFEGEPGPAVSVDTCSSSLVAL 180
 2666 TGTNGODYALRVHNAAGSTGDFALTTAGSVISGRISYTFEGEPGPAVSVDTCSSSLVAL 180
 QY 181 HLAQALRAGECSMALAGVTWSSPGAFAVEFSRQGLADGCKAPSAADGTGMBGV 240
 DB HLAQALRAGECSMALAGVTWSSPGAFAVEFSRQGLADGCKAPSAADGTGMBGV 240
 2726 HLAQALRAGECSMALAGVTWSSPGAFAVEFSRQGLADGCKAPSAADGTGMBGV 240
 QY 241 GMLVRLSDAHNRGHRVLAIVRGSAVNDGASNGLTAPNGPSOORVIRALANAGISAG 300
 DB GMLVRLSDAHNRGHRVLAIVRGSAVNDGASNGLTAPNGPSOORVIRALANAGISAG 300
 2786 GMLVRLSDAHNRGHRVLAIVRGSAVNDGASNGLTAPNGPSOORVIRALANAGISAG 300
 QY 301 DVDAVEAHGTGTLGDPLEAQAALLATYGQDRAGEGPLWLSVKNVGHTOAAAGVAYIX 360
 DB DVDAVEAHGTGTLGDPLEAQAALLATYGQDRAGEGPLWLSVKNVGHTOAAAGVAYIX 360
 2846 DVDAVEAHGTGTLGDPLEAQAALLATYGQDRAGEGPLWLSVKNVGHTOAAAGVAYIX 360
 QY 361 MVMALRGLPRTLHNDPSPHVMSAGAVOLLLETYPWPGEGRLRAGVSSGVSSTIN 420
 DB MVMALRGLPRTLHNDPSPHVMSAGAVOLLLETYPWPGEGRLRAGVSSGVSSTIN 420
 2906 MVMALRGLPRTLHNDPSPHVMSAGAVOLLLETYPWPGEGRLRAGVSSGVSSTIN 420
 QY 421 AHVILEEAPA 430
 DB AHVILEEAPA 430
 2965 AHVILEEAPA 430

RESULT 14

Q9X393
 ID Q9X393 PRELIMINARY; PRT; 6797 AA.
 AC Q9X393
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Polyketide synthase (PIMS1 protein).
 GN PIMS1.
 OS Streptomyces natalensis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=68242;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99214571; PubMed=10187796;
 RA Aparicio J.F., Colina A.J., Ceballos E., Martin J.F.;
 RT "The biosynthetic gene cluster for the 26-membered ring polyene
 RT macrolide pimaricin. A new polyketide synthase organization encoded by
 RT two subclusters separated by functionalization genes.";
 RL J. Biol. Chem. 274:10133-10139(1999).
 DR EMBL: AJ132222; CAB41041.1; -
 DR EMBL: AJ278573; CAC20931.1; -
 DR HSSP; P25715; IMLA.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR004410; FabD.
 DR InterPro; IPR001899; Gram pos anchor.
 DR InterPro; IPR003880; Ketoacyl-synt.
 DR InterPro; IPR000794; Pantone attach.
 DR InterPro; IPR00408; Reg_chir_condens.
 DR Pfam; PF00698; Acyl_transf; 4.
 DR Pfam; PF00109; ketoacyl-synt; 4.
 DR Pfam; PF02801; ketoacyl-synt; 4.
 DR Pfam; PF00550; pp-binding; 4.
 DR TIGRfams; TIGR00128; fabd; 4.
 DR PROSITE; PS00075; ACP_DOMAIN; 4.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 4.
 DR PROSITE; PS00343; GRAM POS ANCHORING; UNKNOWN_1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 3.
 DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
 KW Phosphopantetheine; Transferase.
 SQ SEQUENCE 6797 AA; 710149 MW; 1D2F85074D050DA CRC64;

Query Match 72.6%; Score 1645.5; DB 2; Length 6797;
 Best Local Similarity 73.5%; Pred. No. 6.2e-84;
 Matches 314; Conservative 46; Mismatches 64; Indels 3; Gaps 2;

QY 1 EPPIAVGACRFGPGVTSADDFMDLISSEDAIGGFPFDRGMDLDTLYDPDPHPTGCT 60
 DB EPIAVGACRFGPGVTSADDFMDLISSEDAIGGFPFDRGMDLDTLYDPDPHPTGCT 60
 3299 EPIAVGACRFGPGVTSADDFMDLISSEDAIGGFPFDRGMDLDTLYDPDPHPTGCT 60
 QY 61 RNGEFLYDAGHPDAEFPGISPREALAMPPOORLLLETAMETIEHAGINPHTLHGTPTGVF 120
 DB RSGGFLEHAGFPAEFPGISPREALAMPPOORLLLETAMETIEHAGINPHTLHGTPTGVF 120
 3359 RSGGFLEHAGFPAEFPGISPREALAMPPOORLLLETAMETIEHAGINPHTLHGTPTGVF 120
 QY 121 TGTNGODYALRVHNAAGSTGDFALTTAGSVISGRISYTFEGEPGPAVSVDTCSSSLVAL 180
 DB TGTNGODYALRVHNAAGSTGDFALTTAGSVISGRISYTFEGEPGPAVSVDTCSSSLVAL 180
 3419 TGTNGODYALRVHNAAGSTGDFALTTAGSVISGRISYTFEGEPGPAVSVDTCSSSLVAL 180
 QY 181 HLAQALRAGECSMALAGVTWSSPGAFAVEFSRQGLADGCKAPSAADGTGMBGV 240
 DB HLAQALRAGECSMALAGVTWSSPGAFAVEFSRQGLADGCKAPSAADGTGMBGV 240
 3477 HLAQALRAGECSMALAGVTWSSPGAFAVEFSRQGLADGCKAPSAADGTGMBGV 240
 QY 241 GMLVRLSDAHNRGHRVLAIVRGSAVNDGASNGLTAPNGPSOORVIRALANAGISAG 300
 DB GMLVRLSDAHNRGHRVLAIVRGSAVNDGASNGLTAPNGPSOORVIRALANAGISAG 300
 3537 GMLVRLSDAHNRGHRVLAIVRGSAVNDGASNGLTAPNGPSOORVIRALANAGISAG 300

QY	301	DVDVAEAGTGTLLCDPIEAQQLLATYQODRAGEBPLMGYSKSVNGTQAAAGYIX	360
Db	3597	DVDVAEAGTGTLLCDPIEAQQLLATYQGBREDRPLLGYSKSIIGHAQAASGAYIX	3656
QY	361	MYMALRHLLPRTLVHDEBSPHVMWSAGAVQLTETVMPQGEGLRAAGYSSFGVSGTN	420
Db	3657	MYLAKRHGLPRTLVHDEBSSHVMWSAGAVELLTSEAMPOGEGP-KRAGVSSFGVSGTN	3715
QY	421	AHVILIE	427
Db	3716	AHVILIO	3722

RESULT 15	
Q93NW6	
ID Q93NW6	PRELIMINARY;
Q93NW6	PRT; 10917 AA

01-DEC-2001	(TREMBLrel. 19, Created)
01-DEC-2001	(TREMBLrel. 19, Last sequence update)
01-JUN-2002	(TREMBLrel. 21, Last annotation update)
AMPHC.	
AMPHC.	
Streptomyces nodosus.	
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;	
Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes	
NCBI_TaxId=40318;	
[1]	
SEQUENCE FROM N.A.	
Caffrey P., Lynch S.V., Flood E.M., Finnan S.M., O'Leary M.;	
"The amphotericin biosynthetic gene cluster from Streptomyces	
nodosus";	
Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.	
EMBL; AF357202; AAU73514.1; -	
InterPro; IPR001227; AC transferase.	
InterPro; IPR002328; ADH zinc.	
InterPro; IPR002085; adh_zn_family.	
InterPro; IPR004410; PabD.	
InterPro; IPR000794; ketoacyl-synt.	
InterPro; IPR003880; Pantane attach.	
Pfam; PF00698; Acyl_transfer.6.	
Pfam; PF00107; adh_zinc.1.	
Pfam; PF00109; ketoacyl-synt.6.	
Pfam; PF02801; ketoacyl-synt_C.6.	
Pfam; PF00550; pp-binding.6.	
TIGRfam; TIGR00128; fabd.6.	
PROSITE; PS50075; ACP DOMAIN.6.	
PROSITE; PS00059; ADH ZINC; UNKNOWN 1.	
PROSITE; PS00606; B KETOACYL SYNTHASE; UNKNOWN 6.	
PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_5.	
phosphopantetheine.	
SEQUENCE 10917 AA; 1132905 MW; 15AC5956BBS810A1 CRC64;	

Query Match	72.5%	Score 1641	DB 2	Length 10917
Best Local Similarity	73.0%	Pred. NO. 2e-83		
Matches 316		Conservative 45	Mismatches 66	Indels 6
				Gaps 3

[illegible]

Db	3739	GMLLEHQSHIRNGHEITLAVKSGAVNQDGSNGLTAPMGPSQQRVIRQALLSGSLGTG	3798
Qy	301	DVDVAEHAHGTTGLGDEIEAQALLATATYGDPRAGEPLMGSVKSNVGHTOAAAGVAYIK	360
Db	3799	DVDVAEHAHGTTGLGDEIEAQALLATATYGNRTPDEQPLMLSLSNIGHTQAAAGVAYIK	3858
Qy	361	MMVALRHGLPRTLHVDEPSPHVDSAGAVOLLLETYPMGEGEGLRLRAGVSSFGVSGTN	420
Db	3859	MMMSMRGCVLPCLTHVAPSSHDMTGEAVELLTQTEWFPETD-RVRRAAVSSFGISGTN	3917
Qy	421	AHVILIEE---APA	430
Db	3918	AHVILIEQPVVAPA	3930

Search completed: June 17, 2003, 13:07:56
Job time : 20.8736 secs

THIS PAGE BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 17, 2003, 12:52:22 ; Search time 4.66929 Seconds
(without alignments)
3803.313 Million cell updates/sec

Title: US-09-914-286-3_COPY_2086_2515

Perfect score: 2265
Sequence: 1 EPIAIVGMACRPGCVTSAD.....VSSFGVSGTNHVLSEAPA 430

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1654	73.0	3519	1	OL56_STRAT
2	1544.5	68.2	3567	1	ERY2_SACER
3	1495.5	66.0	3491	1	ERY1_SACER
4	1425.5	62.9	3172	1	ERY3_SACER
5	1104.5	48.8	1538	1	PPSB_MYCTU
6	1066	47.1	1876	1	PPSA_MYCTU
7	1051	46.4	2110	1	MCAS_MYCBO
8	865.5	38.2	4447	1	PKSL_BACSU
9	843	37.2	4427	1	PKSL_BACSU
10	830	36.6	1774	1	MSAS_PENPA
11	793	35.0	4273	1	PKSM_BACSU
12	713	31.5	2511	1	PAS_CHICK
13	702.5	31.0	2181	1	STCA_EMENT
14	684.5	30.2	2109	1	PKSI_ASPPA
15	682	30.1	2505	1	PAS_FAT
16	681	30.1	1986	1	WA_EMENT
17	674.5	29.8	2504	1	PAS_HUMAN
18	628	27.7	352	1	PAS_ANSAN
19	370.5	16.4	401	1	NOE4_RHIME
20	368	16.2	402	1	NOE4_RHIME
21	365	16.1	426	1	KASI_STRGA
22	354	15.6	402	1	NOE_RHIS3
23	352	15.5	403	1	NOE_RHIV
24	343.5	15.2	424	1	KASA_STRCO
25	342	15.1	421	1	FABF_RHIME
26	337.5	14.9	420	1	NODE_RHILT
27	332	14.7	412	1	FABF_ECOLI
28	325	14.3	420	1	KASI_STRCM
29	319	14.1	417	1	FAB2_MYCTU
30	317.5	14.0	422	1	KASI_STRCN
31	317.5	14.0	422	1	CEMI_YEAST
32	312.5	13.8	422	1	KASI_STRHA
33	308.5	13.6	413	1	FABF_VIBCH

34	307	13.6	414	1	FABF_VIBHA	P55338 vibrio harv
35	304	13.4	421	1	KASI_STRVN	P16540 streptomyc
36	292.5	12.9	462	1	FAB2_HORVU	P23902 hordum vul
37	290.5	12.8	420	1	FAB2_MYCLE	O69473 mycobacteri
38	288	12.7	416	1	FAB1_MYCLE	O96b57 mycobacteri
39	286	12.6	416	1	FAB1_MYCTU	O10524 mycobacteri
40	281.5	12.4	406	1	FAB2_ECOLI	P14926 escherichia
41	280.5	12.4	416	1	FAB2_SYNY3	P73283 synechocyst
42	279	12.3	473	1	FAB2_ARATH	P52410 arabidopsis
43	277.5	12.3	423	1	KASI_STRCO	P23155 streptomyc
44	261.5	11.5	406	1	FAB2_HARIN	P43770 haemophilus
45	256	11.3	416	1	PKSF_BACSU	P40804 bacillus su

ALIGNMENTS

RESULT 1	ID	OL56_STRAT	STANDARD	PRT	3519 AA.
AC	007017				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Oleandomycin	polycyclic ketone, modules 5 and 6.			
GN	ORF.				
OS	Streptomyces antibioticus.				
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.				
OX	NCBI_TaxID=1890;				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94150470; Pubmed=8107683;				
RA	Shan D.G., Rodriguez A.M., Valiches C., Mendez C., Salas J.A.;				
RT	"Characterization of a Streptomyces antibioticus gene encoding a type				
RL	I polycyclic ketone synthase which has an unusual coding sequence."				
CC	Mol. Gen. Genet. 242:358-362(1994).				
CC	-1- FUNCTION: MAY BE INVOLVED IN THE BIOSYNTHESIS OF THE OLEANDOMYCIN				
CC	LACTONE RING.				
CC	-1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.				
CC	-1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.				
CC	-----				
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION				
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTREACH				
CC	THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS				
CC	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY				
CC	MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL				
CC	ENTITIES REQUIRES A LICENSE AGREEMENT (See http://www.isb.ch/announce/				
CC	OR SEND AN EMAIL TO license@isb.ch).				
CC	-----				
DR	EMBL; L09654; AAA19695.1; -.				
DR	HSP; P25715; 1MLA.				
DR	InterPro; IPR001227; Ac transferase.				
DR	InterPro; IPR000794; ketoacyl-synt.				
DR	InterPro; IPR003880; Pantoic attach.				
DR	InterPro; IPR001031; Thioesterase.				
DR	Pfam; PF005109; ketoacyl-synt; 2.				
DR	Pfam; PF00550; pp-binding; 2.				
DR	Pfam; PF00658; Acyl-transf; 2.				
DR	Pfam; PF00975; Thioesterase; 1.				
DR	Pfam; PF02801; ketoacyl-synt C; 2.				
DR	PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.				
DR	PROSITE; PS00056; B KETOACYL SYNTHASE; 2.				
DR	PROSITE; PS50075; ACP_DOMAIN; 2.				
KW	transferase; Acyltransferase; Antibiotic biosynthesis; NADP;				
KW	Phosphopantetheine; Multifunctional enzyme; Repeat.				
FT	DOMAIN	1	3519		MODULE 5.
FT	DOMAIN	?	?		MODULE 6.
FT	DOMAIN	32	501		BETA-KETOACYL SYNTHASE 1.
FT	DOMAIN	565	890		ACTINTRANSFERASE (AT) 1.
FT	DOMAIN	1200	1382		BETA-KETOACYL REDUCTASE 1.
FT	DOMAIN	1487	1561		ACYL CARRIER (ACP) 1.
FT	DOMAIN	1686	2156		BETA-KETOACYL SYNTHASE 2.

FT DOMAIN 2220 2541 ACYLTRANSFERASE (AT) 2.
 FT DOMAIN 2856 3038 BETA-KETOACYL REDUCTASE 2.
 FT DOMAIN 3141 3215 ACYL CARRIER (ACP) 2.
 FT ACT_SITE ? ? ACYL-ENZYME INTERMEDIATE.
 FT BINDING ? ? PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT DOMAIN 3270 3519 THIOESTERASE.
 FT ACT_SITE 210 210 BETA-KETOACYL SYNTHASE.
 FT NP BIND 1203 660 ACYL-ENZYME INTERMEDIATE.
 FT BINDING 1524 1524 NADP.
 FT ACT_SITE 1859 1859 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT ACT_SITE 2311 2311 BETA-KETOACYL SYNTHASE.
 FT NP BIND 2859 2905 ACYL-ENZYME INTERMEDIATE.
 FT BINDING 3178 3178 NADP.
 FT SEQUENCE 3519 AA; 368561 MW; 41AE78AAAB61F86 CRC64;
 SQ
 Query Match 73.0%; Score 1654; DB 1; Length 3519;
 Best Local Similarity 73.3%; Pred. No. 8.8e-106;
 Matches 315; Conservative 40; Mismatches 75; Indels 0; Gaps 0;
 QY 1 EPIAIVGACRPPGVSADDFMFLISEQDAIGFPPDRGMDLDTLYDDPDHRCYCT 60
 DB 1687 EAIITATNSCRFGGIDSPEDLMRFLAEGDAVAGLPEDRGMDLDTLYDDPDHRCYCT 1746
 QY 61 RMGSEFLYDAGHFDAPFGISPREALAMPQORLLLETAMETIEHAGINPHTLGTPTGVF 120
 DB 1747 RBGAFRDADAGFDAGFGISPREALAMPQORLLLETAMETIEHAGINPHTLGTPTGVF 1806
 QY 121 TGTNGQDYALRHNAGOSTDGFALTGTAGSVISGRISTTPEPEGPANVSPTAGSSLYAL 180
 DB 1807 IGACHOGGPDPRKAPESVAGVYLLTGTASAVLSGISTYFGLPEPAVVDVTRACSSLYAL 1866
 QY 181 HLACQALRAGCSNMAAGVTVMSPPGAFVFEFSRORGLAAGCHKAPSAADGQMGEGV 240
 DB 1867 HLAQALRAGCSNMAAGVTVMSPPGAFVFEFSRORGLAAGCHKAPSAADGQMGEGV 1926
 QY 241 GMLVERLSDAHNGHRVLAVERGSAVNOGASNGLTAPNGSOORVLRQALANAGLSAG 300
 DB 1927 SLILLERLSDARRIGHRVLAVERGSAVNOGASNGLTAPNGSOORVLRQALANAGLSAG 1986
 QY 301 DVDAVEAHGTTTGGDPIEAQALLATTYGODPAGSGPLTSGVKSNGVHTQAAAGVAYIK 360
 DB 1987 DVDAVEAHGTTTGGDPIEAQALLATTYGODPAGSGPLTSGVKSNGVHTQAAAGVAYIK 2046
 QY 361 MVMALRGLPRTLHVDESPHVMSAGAVOLLTETVWPGSERLRAGVSGVSGTN 420
 DB 2047 MVMALRGLPRTLHVDESPHVMSAGAVOLLTETVWPGSERLRAGVSGVSGTN 2106
 QY 421 AHVILBEAPA 430
 DB 2107 AHVILBEAPA 2116
 RESULT 2
 ERY2 SACER STANDARD; PRT; 3567 AA.
 ID ERY2 SACER
 AC 003132; 054096; PRT; 3567 AA.
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Erythronolide synthase, modules 3 and 4 (EC 2.3.1.94) (ORF 2) (6-deoxyerythronolide B synthase II) (DEBS 2).
 GN ERY2
 OS Saccharopolyspora erythraea (Streptomyces erythraeus).
 CC Bacteria; Actinobacteriales; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;
 CC Saccharopolyspora.
 OC NCBI_Taxid=1836;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220065; PubMed=2024119;
 RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
 RT "Modular organization of genes required for complex polyketide

RT biosynthesis.";
 RL Science 252:675-679 (1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL 2338;
 RX MEDLINE=912155230; PubMed=1740151;
 RA Bevilacqua J., Cortes J., Haydock S.F., Leadlay P.F.;
 RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme."
 RL Eur. J. Biochem. 244:39-49 (1992).
 CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-deoxyerythronolide B.
 CC -1- COFACTOR: NADP; CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETHEINES.
 CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
 CC -1- MISCELLANEOUS: IN EACH ORF OF ERY2 TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERY2 SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX FAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
 CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR), DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.
 CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M63677; AAA26494.1; -;
 DR EMBL; X62569; CAA44448.1; -;
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR002085; Adh zn family.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; pantane_attach.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl trans; 2.
 DR Pfam; PF02801; ketoacyl-synt C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 DR TRANSFERASE; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KW Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1484
 FT DOMAIN 1485 3567
 FT DOMAIN 27 488
 FT DOMAIN 559 884
 FT DOMAIN 1130 1301
 FT DOMAIN 1397 1467
 FT DOMAIN 1485 1943
 FT DOMAIN 2013 2336
 FT DOMAIN 2383 3066
 FT DOMAIN 3139 3322
 FT DOMAIN 3415 3485
 FT ACT_SITE 202 202
 FT ACT_SITE 651 651
 FT BINDING 1430 1430
 FT ACT_SITE 1661 1661
 FT ACT_SITE 2115 2115
 MODULE 3.
 MODULE 4.
 BETA-KETOACYL SYNTHASE 1.
 ACYLTRANSFERASE (AT) 1.
 BETA-KETOACYL REDUCTASE 1 (POSSIBLY NON-FUNCTIONAL).
 ACYL CARRIER (ACP) 1.
 BETA-KETOACYL SYNTHASE 2.
 ACYLTRANSFERASE (AT) 2.
 DEHYDRATASE/ENOYLREDUCTASE (DH/ER).
 BETA-KETOACYL REDUCTASE 2.
 ACYL CARRIER (ACP) 2.
 THIOESTER BOND.
 ACYL-ENZYME INTERMEDIATE.
 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 THIOESTER BOND.
 ACYL-ENZYME INTERMEDIATE.

FT	NP BIND	2961	2978	NADP (ER).
FT	NP BIND	3142	3157	NADP (ER).
FT	BINDING	3448	3448	PHOSPHOPANETHEINE (BY SIMILARITY).
FT	CONFLICT	438	438	R -> A (IN REF. 2).
FT	CONFLICT	480	480	T -> G (IN REF. 2).
FT	CONFLICT	1241	1241	L -> F (IN REF. 2).
FT	CONFLICT	2664	2664	G -> V (IN REF. 2).
FT	SEQUENCE	3567 AA;	374413 MW;	EE6284F4738AACC0 CRC64;
Query Match 68.2% Score 1544.5; DB 1; Length 3567;				
Best Local Similarity 68.8%; Pred. No. 3.1e-98;				
Matches 296; Conservative 47; Mismatches 84; Indels 3; Gaps 2;				
QY	1	EPIALVGMACRPGGVTSADDPFMDLISSEQDAIGFPPDRGMDLDTLYDPDPDHGTCT	60	
DB	1489	EPIALVIGICRPPGIGSPGEOIMRVLAGAMLTTCGPPADRGWDIGRLYHPDPDNGTSYV	1548	
QY	61	RNGGFLYAGHDPDAEFGGISPREALAMPPOORLLETMETETENAGINPHTLHGPTGVF	120	
DB	1549	DKGGFLTDADDPDPFGITTPREALAMPPOORLLETMETEVAERAGIDPDALRGTDGVF	1608	
QY	121	TGTNGQDYALRVHNAGQSTDPFALTGTAGSVISGRISYTFGFEPAVSVDTACSSSLVAL	180	
DB	1609	VGMNGQSYMQJLLAGEAERVDVGQGLGNASVLSGRITATTFGMEGPAALVDTACSSSLVGI	1666	
QY	181	HLACQALPAGCESMALAGCVTVMSSSPGAFVFEPSRORGLADGHCAPBAADGTGWSGV	240	
DB	1669	HLAMQALRGCECSLTLAGCVTVMSPYTFVDFSTRGALASDGRCAFSARADGPALEGEV	1728	
QY	241	GMLVTERSDAHRNRHRYLAVVRGSAVNDQASNGLTAPNPSQORVIROLANAGLSAG	300	
DB	1729	AALVLEPISRRAPNQHVLAVVRGSAVNDQASNSLAPNPGSDERVIROLANAGVPA	1788	
QY	301	DVDAVEAHGTGTTLGDPIEAQALLATVYQDPRAGEGPLMAGSVKSNVGHGTQAAAGVAYIK	360	
DB	1789	DVDVVEAHGTGTGLGDPPIEAGGLITTYQDR--DRPLRLGSKVTKNIGHTQAAAGVAYIK	1846	
QY	361	MYMALRHGLPRTLHVDEPSPHVDMSAGAVOLLTETVMPGEGRLRRAGVSSFGVSGTN	420	
DB	1847	VVLAMRHGMLPRSLHDELSPHIDWESGAVEVLEEVMPMPAGE--RPRAGVSSFGVSGTN	1905	
QY	421	AHVILEEAPA 430		
DB	1906	AHVILEEAPA 1915		
RESULT 3				
ERY1_SACER STANDARD; PRT; 3491 AA.				
AC	003131;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Erythronolide synthase, modules 1 and 2 (BC 2.3.1.94) (ORF 1) (6-deoxyerythronolide B synthase I) (DEBS 1).			
GN	ERYA.			
OS	Saccharopolyspora erythraea (Streptomyces erythraeus).			
CC	Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;			
CC	Actinomycetales; Pseudonocardiales; Pseudonocardaceae;			
CC	Saccharopolyspora.			
CK	NCBI_TaxID=1836;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9120065; PubMed=2024119;			
RA	Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.,			
RT	"Modular organization of genes required for complex polyketide			
RL	biosynthesis.";			
RL	Science 252:675-679(1991).			
RN	(2)			
RP	SEQUENCE OF 3474-3491 FROM N.A.			
RX	MEDLINE=93231529; PubMed=8386127;			
RA	Donadio S., Staver M.J.,			
RT	"151136, an insertion element in the erythromycin gene cluster of			

RT Saccharopolyspora erythraea.":
 CC Gene 126.147-151 (1993).
 CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
 CC deoxyerythronolide B.
 CC -1- COFACTOR: NADP, CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTETHEINES.
 CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
 CC BIOSYNTHESIS.
 CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH
 CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3
 CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
 CC SYNTHASE PARTICIPATES IN ONE OF THE SIX FAS-LIKE ELONGATION STEPS
 CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
 CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
 CC RESPECTIVELY.
 CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),
 CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER
 CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
 CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
 CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
 CC OF THE FULL-LENGTH CHAIN.
 CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
 CC -1- SIMILARITY: CONTAINS 3 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1db-gib.ch/announce/>
 CC or send an email to license@1db-gib.ch).
 CC -----
 CC EMBL, M63676, AAA26493.2; -.
 CC EMBL, L07626; AAA26504.1; -.
 CC DR HSSP, P25715; IMLA.
 CC DR Interpro: IPR002198; ADH_short.
 CC DR Interpro: IPR001227; Ac_Transferase.
 CC DR Interpro: IPR000794; Ketoacyl-synt.
 CC DR Interpro: IPR003880; Pantne_attach.
 CC DR Pfam, PF00106; adh_short; 1.
 CC DR Pfam, PF00109; ketoacyl-synt; 2.
 CC DR Pfam, PF00550; pp-binding; 3.
 CC DR Pfam, PF00698; Acyl_transf; 3.
 CC DR Pfam, PF02801; ketoacyl-synt_C; 2.
 CC DR PROSITE, PS00012; PHOSPHOPANTETHEINE, 3.
 CC DR PROSITE, PS00606; B_KETOACYL_SYNTHASE, 2.
 CC DR PROSITE, PS50075; ACP_DOMAIN, 3.
 CC KW Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 CC KM Phosphopantetheine; Multifunctional enzyme.
 CC FT DOMAIN 1 1972 MODULE 1.
 CC FT DOMAIN 1979 3491 MODULE 2.
 CC FT DOMAIN 1 375 ACYLTRANSFERASE (AT) 1.
 CC FT DOMAIN 414 484 ACYL CARRIER (ACP) 1.
 CC FT DOMAIN 503 961 BETA-KETOACYL SYNTHASE 1.
 CC FT DOMAIN 1030 1356 ACYLTRANSFERASE (AT) 2.
 CC FT DOMAIN 1611 1794 BETA-KETOACYL REDUCTASE 1.
 CC FT DOMAIN 1888 1958 ACYL CARRIER (ACP) 2.
 CC FT DOMAIN 1979 2441 BETA-KETOACYL SYNTHASE 2.
 CC FT DOMAIN 1979 2854 ACYLTRANSFERASE (AT) 3.
 CC FT DOMAIN 3055 3237 BETA-KETOACYL SYNTHASE 3.
 CC FT DOMAIN 3055 3237 BETA-KETOACYL REDUCTASE 2.
 CC FT DOMAIN 3334 3404 ACYL CARRIER (ACP) 3.
 CC FT ACT_SITE 145 145 ACYL-BINDING INTERMEDIATE.
 CC FT ACT_SITE 447 447 ACYL-BINDING INTERMEDIATE.
 CC FT BINDING 677 677 PHOSPHOPANTETHEINE (BY SIMILARITY).
 CC FT ACT_SITE 1128 1128 BETA-KETOACYL SYNTHASE.
 CC FT NP_BIND 1614 1660 BETA-KETOACYL SYNTHASE.
 CC FT BINDING 1921 1921 ACYL-ENZYM INTERMEDIATE.
 CC FT ACT_SITE 2148 2148 PHOSPHOPANTETHEINE (BY SIMILARITY).
 CC FT BINDING 2598 2598 BETA-KETOACYL SYNTHASE.
 CC FT NP_BIND 3058 3104 ACYL-ENZYM INTERMEDIATE.
 CC FT BINDING 3367 3367 PHOSPHOPANTETHEINE (BY SIMILARITY).
 CC SEQUENCE 3491 AA; 36502 MW; 662BFC32C90FA8C4 CRC64;
 Query Match 66.0%; Score 1495.5; DB 1; Length 3491;

Best Local Similarity 66.2%; Pred. No. 7e-95; Matches 284; Conservative 54; Mismatches 88; Indels 3; Gaps 2;

QY 1 EPVAVVMAACLPFGVSTPBEFWEILSEGRDAVAGLPTRDGMDDSLFHPDPTSGTAHQ 60
 Db 505 EPVAVVMAACLPFGVSTPBEFWEILSEGRDAVAGLPTRDGMDDSLFHPDPTSGTAHQ 564
 QY 61 RANGFLVDAGHFDPAEFGISPREALAMPDPOORLLEETMETIEHAGINPTLHGTPGVF 120
 Db 565 RGGGFLVETATPDAFPGMSPREALAMPDPOORLLEETMETIEHAGINPTLHGTPGVF 624
 QY 121 TGTNGQDYALRVNNAAGSTDDGFPALTGTAGSVISGRISYTFPGFEGPAAVSDTACSSSLVAL 180
 Db 625 VGLIPQEXGPRILACGEGVEGYLMTGTTTSAVSGRIATVLCLEGPALISVDTACSSSLVAL 684
 QY 181 HLAQALPAGCSCMALAGVTVMSSPCAFVFEFSRQGLADGHCKAKSAADGTGMEGV 240
 Db 685 HLAQALPAGCSCMALAGVTVMSSPCAFVFEFSRQGLADGHCKAKSAADGTGMEGV 744
 QY 241 GMLLVERLSDAHRNGHRYLAVRGSAVNOGASNGLTAPNGPSOQRYVROLANAGLSAG 300
 Db 745 GMLLVERLSDAHRNGHRYLAVRGSAVNOGASNGLTAPNGPSOQRYVROLANAGLSAG 804
 QY 301 DVDAVEAHGTGTTGDPLEAQAALLATYGGDRAGEGPIMLGSKVNGHTQAAGVAVIK 360
 Db 805 DIDAVEAHGTGTTGDPLEAQAALLATYGGDRAGEGPIMLGSKVNGHTQAAGVAVIK 862
 QY 361 MWMLARGLPRTLHNDPSPHVMSAGAVOLLLETYPMPGEGRLRAGVSSGVCSTN 420
 Db 863 MWMLARGLPRTLHNDPSPHVMSAGAVOLLLETYPMPGEGRLRAGVSSGVCSTN 921
 QY 421 AHVILEAP 429
 Db 922 AHVILEAP 930

RESULT 4
 ERY3_SACER STANDARD: PRT: 3172 AA.
 AC Q01133; Q54097; Q99270;

DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Erythronolide synthase, modules 5 and 6 (EC 2.3.1.94) (ORF 3) (6-deoxyerythronolide B synthase III) (DEBS 3).
 GN ERYA.
 OS Saccharopolyspora erythraea (Streptomyces erythraeus).
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;
 OC Saccharopolyspora.
 CX NCBI_TaxID=1836;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL 2338;
 RX MEDLINE=91043075; PubMed=2234082;
 RA Cortes J., Haydock S.F., Roberts G.A., Bevilacqua D.J., Leadlay P.F.;
 RT "An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea";
 RL Nature 348:176-178 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9120065; PubMed=2024119;
 RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
 RT "Modular organization of genes required for complex polyketide biosynthesis";
 RL Science 252:675-679 (1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL 2338;
 RX MEDLINE=92155230; PubMed=1740151;
 RA Bevilacqua D.J., Cortes J., Haydock S.F., Leadlay P.F.;
 RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of the structural gene, sequence analysis and inferred domain

RT structure of the multifunctional enzyme.";
 RL Eur. J. Biochem. 204:39-49 (1992).
 CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-deoxyerythronolide B.
 CC -1- COFACTOR: NADP. CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.
 CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
 CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
 CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR), DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.
 CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X56107; CAA39583.1; -;
 DR EMBL: M63677; AAA26495.1; -;
 DR EMBL: X62569; CAA44449.1; -;
 DR HSHP, P00101; 1ICCH.
 DR InterPro: IPR002198; ADH, short.
 DR InterPro: IPR001227; AC, transferase.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantone attach.
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: P00106; adh, short; 1.
 DR Pfam: P00109; ketoacyl-synt; 2.
 DR Pfam: P00550; pp-binding; 2.
 DR Pfam: P00698; Acyl trans; 2.
 DR Pfam: P00975; Thioesterase; 1.
 DR Pfam: P02801; ketoacyl-synt C; 2.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 2.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE: PS50075; ACP DOMAIN; 2.
 KW Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KW Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1484
 FT DOMAIN 1485 3172
 FT DOMAIN 37 484
 FT DOMAIN 554 878
 FT DOMAIN 1116 1298
 FT DOMAIN 1394 1464
 FT DOMAIN 1486 1954
 FT DOMAIN 2021 2335
 FT DOMAIN 2555 2735
 FT DOMAIN 2821 2891
 FT DOMAIN 2926 3172
 FT ACT_SITE 199 199
 FT ACT_SITE 643 643
 FT NP_BIND 1118 1164
 FT BINDING 1427 1427
 FT ACT_SITE 1661 1661
 FT ACT_SITE 2112 2112
 FT NP_BIND 2557 2605
 FT BINDING 2854 2854
 FT CONFLICT 231 231
 FT CONFLICT 240 240
 FT CONFLICT 289 289
 MODULE 5.
 MODULE 6.
 BETA-KETOACYL SYNTHASE 1.
 ACYLTRANSFERASE (AT) 1.
 BETA-KETOACYL REDUCTASE 1.
 ACYL CARRIER (ACP) 1.
 BETA-KETOACYL SYNTHASE 2.
 ACYLTRANSFERASE (AT) 2.
 BETA-KETOACYL REDUCTASE 2.
 ACYL CARRIER (ACP) 2.
 THIOESTERASE.
 ACTL-ENZYME INTERMEDIATE.
 NADP.
 PHOSPHOPANTETHEINE (BY SIMILARITY).
 THIOESTER BOND.
 ACTL-ENZYME INTERMEDIATE.
 NADP.
 PHOSPHOPANTETHEINE (BY SIMILARITY).
 S -> SA (IN REF. 2).
 MISSING (IN REF. 2).
 A -> R (IN REF. 2).

```

FT CONFLICT 493 493 P -> R (IN REF. 2).
FT CONFLICT 493 517 PEPNLSLDTGFTLATRASAMENRA -> ASRGTRCATPVS
FT CONFLICT 510 510 A -> R (IN REF. 2).
FT CONFLICT 513 513 M -> W (IN REF. 2).
FT CONFLICT 525 525 E -> D (IN REF. 2).
FT CONFLICT 536 536 R -> G (IN REF. 2).
FT CONFLICT 547 551 GPNRP -> ARTR (IN REF. 2).
FT CONFLICT 553 553 R -> G (IN REF. 2).
FT CONFLICT 673 673 R -> A (IN REF. 2).
FT CONFLICT 716 716 MISSING (IN REF. 2).
FT CONFLICT 734 736 AHK -> GIT (IN REF. 2).
FT CONFLICT 896 896 R -> RQR (IN REF. 2).
FT CONFLICT 896 896 R -> RELVYPPQRR (IN REF. 1).
FT CONFLICT 988 994 GVAAPH -> VSLSRD (IN REF. 2).
FT CONFLICT 1108 1116 RTHLEPLA -> ARTNSPR (IN REF. 2).
FT CONFLICT 1124 1126 MISSING (IN REF. 1).
FT CONFLICT 1132 1132 L -> V (IN REF. 1).
FT CONFLICT 1192 1192 A -> R (IN REF. 2).
FT CONFLICT 1194 1194 MISSING (IN REF. 2).
FT CONFLICT 1278 1278 AA -> RR (IN REF. 2).
FT CONFLICT 1385 1390 LCGRE -> STARR (IN REF. 2).
FT CONFLICT 1485 1485 MISSING (IN REF. 2).
FT CONFLICT 1518 1518 G -> R (IN REF. 2).
FT CONFLICT 1601 1601 V -> L (IN REF. 2).
FT CONFLICT 1724 1725 LP -> PA (IN REF. 2).
FT CONFLICT 1732 1732 O -> L (IN REF. 2).
FT CONFLICT 1739 1743 GPAEG -> ABRA (IN REF. 2).
FT CONFLICT 1762 1762 T -> S (IN REF. 2).
FT CONFLICT 2252 2252 D -> DGD (IN REF. 2).
FT CONFLICT 2277 2277 QSP -> AVA (IN REF. 2).
FT CONFLICT 2408 2408 G -> GR (IN REF. 2).
FT CONFLICT 2420 2421 LA -> S (IN REF. 2).
FT CONFLICT 2443 2444 NA -> TH (IN REF. 2).
FT CONFLICT 2596 2596 A -> G (IN REF. 2).
FT CONFLICT 2609 2609 P -> A (IN REF. 2).
FT CONFLICT 2715 2722 RRAGRA -> ARKVR (IN REF. 1).
FT CONFLICT 2754 2754 D -> E (IN REF. 2).
SQ SEQUENCE 3172 AA, 331474 MW, DBBD5094E77DD5F CRC64;

```

```

Query Match 62.94; Score 1425.5; DB 1; Length 3172;
Best Local Similarity 65.34; Pred. No. 4.1e-90;
Matches 280; Conservative 44; Mismatches 92; Indels 13; Gaps 3;

QY 1 EPIAIVGACFPFGTYSADPFMDLISEODAGFPDRGMDLTYDPDPHGTCT 60
DB 39 EPIAIVGACFPFGDVDSPESEFMSGGDALNEAPADKGM-----EPDP-----A 86

QY 61 RNGGFLYDAGHFAEFGISPREALAMPQORLLLETAMETIEHAGINPHTLGTGVF 120
DB 87 RLGGMLAAGDPDAGFGISPREALAMPQORIMLEISWEALERAHGDHVSILRSATGVF 146

QY 121 TGTNGOYALRVNAGSGTSGFALTGTNAGYISGRISTTFEFGPNAVSVTACSSLYAL 180
DB 147 TGVGTIVYGPDPAPDEVLGYVTGTASVASVRVAYCGLLEBPATVTDACSSGLTAL 206

QY 181 HLAICALRAGCSMALAGVTWSSPGAFFVFSRORGLAADGHKAPSAADGCMWEGV 240
DB 207 HLAIESLRDEBCGALLAGGVWSSPGAFFVFSRORGLAADGHKAPSAADGCMWEGV 266

QY 241 GMLVBLSDAHRNGHVLAVVRGSAVNODASNGLTAPNGSPQOYVIRQALANAGLSAG 300
DB 267 GVLVLQSLAARREGRPVLAVALAGSAYVNODASNGLTAPNGSPQOYVIRQALANAGLSAG 326

QY 301 DVDAVEAHGCTTIGDPIEAQALLATYGQDAGGCPMLGSKVKNVNGHTQAAAGVAVIK 360
DB 327 DVDAVEAHGCTTIGDPIEAQALLATYGQDAGGCPMLGSKVKNVNGHTQAAAGVAVIK 386

QY 361 MVMALRGLLPTLTHVDEPSPHVMWSAGAVOLLTETVPMPGEGRLRPRAGSVFGVSGTN 420
DB 387 AVLAIRGEMPRTHLHDEPSPQIEMDLAGVSVSQAASWPAFG-RPRAGVSSFGISGTN 445

QY 421 AHVILEAP 429

```

```

DB 446 AHVILEAP 454
|||||
PSSB_MYCTU STANDARD; PRT; 1538 AA.
ID PSSB_MYCTU Q10978; O53234;
DT 01-OCT-1996 (Ref. 34, Created)
DT 30-MAY-2000 (Ref. 39, Last sequence update)
DT 16-OCT-2001 (Ref. 40, Last annotation update)
DE Phenolipidic acid synthesis polypeptide synthase pssb.
GN PSSB OR RV2932 OR MT3002 OR MTCY338.21 OR MTV011.01.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broech R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gass S., Barry C.E. III, Tekala F.,
RA Badcock K., Baeham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagsels K., Krogh A., McLean L., Moul S., Murphy L.,
RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Biegel W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLIPIDIC ACID
CC SYNTHESIS.
CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETHEINE (BY
CC SIMILARITY). CONTAINS 1 ACYL CARRIER DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL; AL021070; CA15929.1; -
DB EMBL; AE007122; AA47329.1; -
DB TIGR; MT3002; -
DB Tuberculist; RV2932; -
DB InterPro; IPR001227; Ac transferase.
DB InterPro; IPR000794; Ketaoacyl-synt.
DB InterPro; IPR003880; Pantne_attach.
DB Pfam; PF00109; Ketaoacyl-synt; 1.
DB Pfam; PF00698; Acyl transferase; 1.
DB Pfam; PF02801; Ketaoacyl-synt; 1.
DB PROSITE; PS00606; B KETOACYL SYNTHASE; 1.
DB PROSITE; PS00012; PHOSPHOPANTHETHEINE; FALSE_NEG.
DB PROSITE; PS50075; ACP DOMAIN; 1.
KW Hypothetical protein; Multifunctional enzyme; Oxidoreductase;
KW transferase; NADP; Phosphopantetheine; Complete proteome.

```

FT DOMAIN 1425 1495 ACYL CARRIER (ACP).
 FT BINDING 1458 1458 PHOSPHOPANTETHEINE (POTENTIAL).
 SQ SEQUENCE 1538 AA; 162527 MW; B5582A042AD00CC CRC64;

Query Match 48.8%; Score 1104.5; DB 1; Length 1538;
 Best Local Similarity 52.5%; Pred. No. 2.2e-68;
 Matches 224; Conservative 64; Mismatches 134; Indels 5; Gaps 3;

1 EPIAIVGACRFPGVTSADDFMDLISSEDAIGFPPDRGMDLTLYPDPPHPTGCT 60
 DB 34 EPVAVVIGCRFPFGVDPGSPFSPFVAGRNALSTVPADR-WDAEATFHPDPLTGRMTT 92
 QY 61 RNGFLYDAGHFAEFPGISPREALAMPDQORLLLETAMETIEHAGINPHTLHGTPGVF 120
 DB 93 KMGGFVDPVAGFDLAEFGITPREAAMPDQORMLLEVAMELEHAGIPDGLGTRAVM 152
 QY 121 TGTNGQDYALRVHNAAGSTDFALTGTAGSVISGRISTTFPEEGPANSVDTACSSSLVAL 180
 DB 153 MGVEFNEYSMLASPOVVDVAVSGTGNASHITVGRISVTLGLRGPVAVADVACSSSLVAV 212
 QY 161 HLACQALRAGECSMALAGVTVMSPGAFFVEFSRQGLAANDHCKAFSAADGTGMEGV 240
 DB 213 HLACQSLRLRETDALAGGVSTTRPETQIAISMGSLPQGRCAAFDAADGFRVREGGA 272
 QY 241 GMLVERLSDAHRNGHRYLAVVRSAGVNOGASNGLTAPNGPSQORVTRQALANAGLSAG 300
 DB 273 GVAVLAKRLTDVAVRGDDVLAVRSSAVVQDGRSNGVAPNPAACDVADALRSGDVAD 332
 QY 301 DVDAVEAHGTGTTGDPLEAALLATYGQDPAGSGPLMSGVKSNGVHTQAAAGVAVIK 360
 DB 333 SVNVVEAHGTGTGVDPLIEFALAAATYGH--GGDACALGAVKNIGLEMAAGIAGTIK 389
 QY 361 MMVALRGLPLRTLAVDESPHVMWSAGAVOLLTETVWPFGEGRLRAGVSSGVSSTN 420
 DB 390 ATLAVQARTIPNHLHFSQMNPAIDAASFRFVPTQNSFWPFAEGP-RAAVSSFGLGSTN 448
 QY 421 AHVILEE 427
 DB 449 AHVILEQ 455

RESULT 6
 ID_PPSA MYCTU STANDARD; PRT; 1876 AA.

AC 010977;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenolphthalein synthesis polypeptide synthase ppsa.
 GN PPSA OR RV2931 OR MT3000 OR MTCY338.20.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_taxid=1773;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; Pubmed=96343230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigemeier K., Gao S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Davlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).

RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Ohkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Ustebach T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.,
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POLYPEPTIDE SYNTHASE INVOLVED IN PHENOLPHTHALEIN
 CC SYNTHESIS.
 CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).

DR EMBL; Z74697; CAA98988.1; -;
 DR EMBL; AB007122; AAK47328.1; -;
 DR TIGR; MT3000; -;
 DR TubercuList; RV2931; -;
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Phantne_attach.
 DR Pfam; PF00109; ketoacyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl_transf; 1.
 DR Pfam; PF02801; ketoacyl-synt C; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; 1.
 DR PROSITE; PS00075; ACP DOMAIN; 2.
 DR KX Multifunctional enzyme; Oxidoreductase; Transferase; NADP; Repeat;
 KM Phosphopantetheine; Complete proteome.

FT DOMAIN 7 80
 FT NP_BIND 1764 1833
 FT BINDING 62 62
 FT ACT_SITE 273 273
 FT ACT_SITE 720 720
 FT BINDING 1796 1796
 FT CONFLICT 624 624
 FT CONFLICT 877 877
 FT CONFLICT 1323 1323
 SQ SEQUENCE 1876 AA; 198834 MW; D9783DBD48792110 CRC64;

Query Match 47.1%; Score 1066; DB 1; Length 1876;
 Best Local Similarity 50.0%; Pred. No. 1.2e-65;
 Matches 215; Conservative 66; Mismatches 141; Indels 8; Gaps 4;

1 EPIAIVGACRFPGVTSADDFMDLISSEDAIGFPPDRGMDLTLYPDPPHPTGCT 59
 DB 102 EPIAIVGACRFPGVTSADDFMDLISSEDAIGFPPDRGMDLTLYPDPPHPTGCT 157
 QY 60 --TRNGFLYDAGHFAEFPGISPREALAMPDQORLLLETAMETIEHAGINPHTLHGTP 117
 DB 158 RTRKMSGFLPDIDAFDAEFPEISPSSEADKNDPQORLLLEVAMELEHAGIPDGLGTRAT 217
 QY 118 GVFTGTNGQDYALRVHNAAGSTDFALTGTAGSVISGRISTTFPEEGPANSVDTACSSSL 177
 DB 218 GVFAAGLSTYGMASADLSQVDGMSNGAMITIANRLSYFLDLKGPSAVAVTACSSSL 277
 QY 178 VALHACQALRAGECSMALAGVTVMSPGAFFVEFSRQGLAANDHCKAFSAADGTGMEGV 237
 DB 278 VALHACQSLRLRETDALAGGVSTTRPETQIAISMGSLPQGRCAAFDAADGFRVREGGA 337
 QY 238 EGVMGLVERLSDAHRNGHRYLAVVRSAGVNOGASNGLTAPNGPSQORVTRQALANAGL 297
 DB 338 EGAGVAVLAKRLTDVAVRGDDVLAVRSSAVVQDGRSNGVAPNPAACDVADALRSGDVAD 397

QY 298 SAGDVDAVEAHGCTTLLGDPLEAQAALLATYGGODRAGEGRLWLSGKSNVGHGTOAAAGVAG 357
 DB 398 QPSEVDVEAHGCTTLLGDPLEAQAALLATYGGODRAGEGRLWLSGKSNVGHGTOAAAGVAG 457
 QY 358 VIKVMALRLHGLPRTLHDEPSPHVMASAGAVOLLTETYPWPGEGRLRACGSSGCVS 417
 DB 458 FIKTVLAVQHQGPDPNQHPTANPHIPFTDLRMKVDTQTEWP-ATGHPRAVAGVSSFGG 516
 QY 418 GTNAHVILE 427
 DB 517 GTNAHVILEQ 526
 RESULT 7
 MCAS MYCBO STANDARD; PRT, 2110 AA.
 ID MCAS MYCBO
 AC 002251;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Mycocerosic acid synthase.
 GN MAS.
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 ON NCBI_TaxID=1765;
 RX SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP MEDLINE=92406887; PubMed=1527058;
 RA Mehnert M., Kolatukudy P.E.;
 RT "Molecular cloning and sequencing of the gene for mycocerosic acid
 synthase, a novel fatty acid elongating multifunctional enzyme, from
 Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guérin.";
 RT J. Biol. Chem. 267:19388-19395(1992).
 CC -1- FUNCTION: CATALYZES THE ELONGATION OF N-PARTY ACYL-COA WITH
 METHYLMALONYL-COA (NOT MALONYL-COA) AS THE ELONGATING AGENT TO
 FORM MYCOCEOSYL LIPIDS.
 CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETINE.
 CC -1- SUBUNIT: HOMODIMER WHOSE MONOMERS PROBABLY HAVE A HEAD TO TAIL
 ARRANGEMENT.
 CC -1- SUBCELLULAR LOCATION: Membrane-associated.
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M95808; AAA25369.1; -.
 DR PIR; B44110; B44110.
 DR HSP; P73283; 1ESM.
 DR InterPro; IPR001227; AC transferase.
 DR InterPro; IPR002085; Adh zn family.
 DR InterPro; IPR000794; Ketocyl-synt.
 DR InterPro; IPR003880; Ppancne_attach.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketocyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR Pfam; PF00698; ACYL transfer; 1.
 DR Pfam; PF02801; ketocyl-synt C; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE; PS50075; ACP_DOMAIN; 1.
 KM Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 KM Transferase; Hydroxylase; Oxidoreductase; Ligase; NADP; Membrane.
 FT DOMAIN 1 430 BETA-KETOACYL SYNTHASE.
 FT DOMAIN 533 852 ACYL TRANSFERASE.
 FT DOMAIN ? ? ENOYL REDUCTASE.
 FT DOMAIN ? ? BETA-KETOACYL REDUCTASE.
 FT DOMAIN 2026 2096 ACYL CARRIER (ACP).

FT ACT SITE 177 177 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT SITE 623 623 ACYL TRANSFERASES (BY SIMILARITY).
 FT NP_BIND 1561 1578 NADP (ER).
 FT NP_BIND 1765 1780 NADP (KR).
 FT BINDING 2059 2059 PHOSPHOPANTHETINE (BY SIMILARITY).
 SQ SEQUENCE 2110 AA; 225577 MW; CC658215D7155300 CRC64;
 Query Match 46.4%; Score 1051; DB 1; Length 2110;
 Best Local Similarity 50.3%; Pred. No. 1.5e-64;
 Matches 218; Conservative 61; Mismatches 142; Indels 12; Gaps 7;
 QY 2 PLAIYMACRFGCVTSADDFMDLISSEODALGCFPTDRGMDLDTLYDPDPHGTCTYR 61
 DB 7 PVAVIGWGCRLPGGINSPKLWESLRLGCDLVTETPPDR-WDADPYDEPVGGRSVSR 65
 QY 62 NGGELVDAGHPDPEFEFGISPREALMDPOORLLTFAMETIEHAGINPHLTGCTPGVFT 121
 DB 66 WGGFLDVAQFDPEFFGISERENTSIDPOORLLTETSWAIEHAGLDPAISLAVFT 125
 QY 122 GTNGQDYALRVHNAQSTGDFALGTAGSVISGRISYTFGEGPAVSVDTCSSVALH 181
 DB 126 GLTHEDYLVLTITAGGLASPVVVTGLANSVASGRINHTGLHGPAMTPTDAGSGMAVA 185
 QY 182 LACQALRAGECSMALAGVTVNWSPPGAFVEFSRQGLADGHCASFSAADGTGWEQVG 241
 DB 186 LACRSIHDERDLALAGCAVLLEPHACVAASQGMLSSTGRCHSFADADGFRSEGCA 245
 QY 242 MLVLELSDAHRGHVLAIVGSAVNOGASNGLTAPNGSPQRIYRQALNAGSAGD 301
 DB 246 MVLKRLPALRQGNIFAVVRGTATNODGRTETLTMPBEDAVAAVRAALAAVQVPT 305
 QY 302 VDAVEAHGCTTLLGDPLEAQAALLATYGGODRAGEGRLWLSGKSNVGHGTOAAAGVAG 361
 DB 306 VGVVEAHGCTPTGDIETRSRLARYG--AGI-PALGASNSMGSHTASAGTVGLIRA 361
 QY 362 VVALRHGLPRTLHDE-PSPHVMSAG--AVQLTETYPWPGEGRL-RRACVSSFGVS 417
 DB 362 ILSLRGVVPLHLFRLLDELSDVETGLFVPAVT---PWPNGNHTPRKRVAVSSFGVS 418
 QY 418 GTNAHVILEAPA 430
 DB 419 GTNAHVILEAPA 431
 RESULT 8
 PKSK_BACSU STANDARD; PRT, 4447 AA.
 ID PKSK_BACSU
 AC P40803;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative polyketide synthase pksK (PKS).
 GN PKSK.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / PB1424;
 RA MEDLINE=95219083; PubMed=7704258;
 RA Albertini A.M., Carimori T., Scofield F., Scotti C., Gallazzi A.;
 RT "Sequence around the 159 degree region of the Bacillus subtilis
 genome: the pksK locus spans 33.6 kb.";
 RT Microbiology 141:299-309(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunet F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
 RA Borries R., Bortner L., Brans A., Braun M., Briganti S.C., Bron S.,
 RA Brouillet S., Bruchet C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.D., Connerton I.F., Cumming N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Ertian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
 RA Ghm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Hernat A.,
 RA Hilbert H., Holappell S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaier-Blanchard M., Klein C.,
 RA Kobayashi Y., Koelter P., Konigstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,
 RA Medina N., Melillo R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Paro V., Pohl T.M., Portetelle D., Porvolik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter K., Scottone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Solito B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vaasaroti A.,
 RA Viari A., Wandut R., Weder E., Weder H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumestein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*,"
 RT Nature 390:249-256(1997).
 RL -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR
 CC THE SYNTHESIS OF A POLYPEPTIDE MOLECULE WHICH MAY BE INVOLVED IN
 CC SECONDARY METABOLISM.
 CC -1- COPACATOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES
 CC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 5 ACYL CARRIER DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U1039; AAA85144.1; -;
 DR EMBL: Z99112; CAB13590.1; -;
 DR EMBL: Z99113; CAB13601.1; -;
 DR HSSP: P14687; IAMU.
 DR Subtilist; BGI0930; pkak.
 DR InterPro; IPR000873; AMP-bind.
 DR InterPro; IPR001242; Condensacn.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantine_attach.
 DR Pfam; PF00109; ketoacyl-synt; 3.
 DR Pfam; PF00501; AMP-binding; 1.
 DR Pfam; PF00550; pp-binding; 5.
 DR Pfam; PF00668; Condensation; 1.
 DR Pfam; PF02801; ketoacyl-synt_C; 3.
 DR PRINTS; PR00154; AMPBINDING.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
 DR PROSITE; PS00455; AMP BINDING; 1.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
 DR PROSITE; PS00075; ACP DOMAIN; 5.
 KM Transferase; Acyltransferase; Antibiotic biosynthesis; NADP;
 KM Phosphopantetheine; Multifunctional enzyme; Repeat; Ligase;
 KM Complete proteome.
 FT DOMAIN 1 68 ACYL CARRIER (ACP) 1.
 FT DOMAIN 1063 1130 ACYL CARRIER (ACP) 2.
 FT DOMAIN 2516 2589 ACYL CARRIER (ACP) 3.
 FT DOMAIN 2618 2687 ACYL CARRIER (ACP) 4.
 FT DOMAIN 3688 3937 ACYL CARRIER (ACP) 5.
 FT BINDING 31 31 PHOSPHOPANTETHEINE (POTENTIAL).
 FT BINDING 1093 1093 PHOSPHOPANTETHEINE (POTENTIAL).
 FT BINDING 2552 2552 PHOSPHOPANTETHEINE (POTENTIAL).

FT BINDING 2650 2650 PHOSPHOPANTETHEINE (POTENTIAL).
 FT ACT SITE 2915 2915 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT BINDING 3900 3900 PHOSPHOPANTETHEINE (POTENTIAL).
 FT ACT SITE 4147 4147 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 SQ SEQUENCE 4447 AA; 496058 MW; 9C6DBA46C0A0C057 CXC64;
 Query Match 38.2%; Score 865.5; DB 1; Length 4447;
 Best Local Similarity 41.8%; Pred. No. 2, 1e-51;
 Matches 185; Conservative 80; Mismatches 159; Indels 19; Gaps 8;
 QY 1 EPIAIVGMAACRPFQGVTSADDFWDLISSEDAIGFPTDRGMDLTLV-DPPDHETCY 59
 DB 2744 EPPAIVGISGRFPGAM-DIDFENKLEBKDSITEVPKOR-WDMREHYGNPDVVKT-D 2800
 QY 60 TRNGGFLYDGHDPDAEPFGISPREALAMPQOGLLETMETTEHAGINHTLHPTGV 119
 DB 2801 IKMGGFLDVAEPDPLFPGISPREAYVDPOQLMTYVKALEDGSCQSLSGTGTG1 2860
 QY 120 FTGNGQDYALRHNNAGQSTDFALTG-TAGVISGRISYTFEGEPRAVSDTRACSSSLV 178
 DB 2861 FICTGNTGYDLPFRANLPIEGHAATGHMIPSGPRMSTFLNHGSPREVEATACSSSLV 2920
 QY 179 ALHACQALPAGECSMALAGCTVWSSPGAFVFSRQGLADGHCKAFAAADGTGWE 238
 DB 2921 AIHRAVTAQNGDCEMAIAGVNTIITEAHISYSKAGMISTGRCKTFSDANGYVGE 2980
 QY 239 GUGMLVERLSDAHNRCHRLAVVRSAYVQDASNGTLTPNPSQORVROLANAGLS 298
 DB 2981 GVGWVLMKLEDERGDNHNYGVIRGAENHGRANTLSPNPAQDLVLRVRYROADID 3040
 QY 299 AGVDVAEAGTGTTLGDPPIEAQALLATYGO--DRAGEGP-----LMLGYSVNVGHT 349
 DB 3041 PSTVITYEANGTGTLELGPPIEINGLKAFELSNMGESGPDVDRHRCIGSVKSNIGHL 3100
 QY 350 QAAAGVAGVIRKVMALRHGLPRTLHVDEPSPHVMSAGAVQLITETVWPG-----GEG 404
 DB 3101 ELAAGISGLIKVLLQMKHKLTVLSHCETINPLYQLTDSFPIYOEKQEWKSVTDHROGNE 3160
 QY 405 RLRRAGVSSRSGVGTNAHVLEE 427
 DB 3161 LPRRAGISSFGIGGVNNAHVIEE 3183
 RESULT 9
 PKSL_BACSU STANDARD; PRT; 4427 AA.
 ID PKSL_BACSU
 AC Q05470;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative polyketide synthase pkel (PKS).
 GN PKSL OR PKSX OR PKSA OR OUTG.
 OS *Bacillus subtilis*.
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / PB1424;
 RX MEDLINE=9345824; PubMed=8344529;
 RA Scotti C., Piatelli M., Cuzzoni A., Perani P., Tognoni A., Grandi G.,
 RA Galizzi A., Albertini A.M.;
 RT "A *Bacillus subtilis* large ORF coding for a polypeptide highly
 RT similar to polyketide synthases.";
 RL Gene 130:65-71(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Betero M.G., Beesieres P., Bolotin A., Borchett S.,
 RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brunsch C.V., Caldwell B., Capiano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

RA Dentsoe F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Ertlen K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gilm S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holstappel S., Hosono S., Hullo M.P., Itaya M., Jones L.,
RA Jorle B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapides A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Maeda S., Manuel C., Medigne C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nodack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Potwollik S., Prescott A.M.,
RA Pesecean E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche M., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Solde B.,
RA Sorokin A., Taccoul E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tobeo V., Uchiyama S., Vandenbol M., Vannier P., Vassaret A.,
RA Viari A., Wambert R., Wedler E., Wedler H., Weitzenger T.,
RA Winters P., Wipert A., Yamamoto H., Yamane K., Yasauno K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis." ;
RL Nature 390:249-256(1997).
RN [3]
RP SEQUENCE OF 3619-4427 FROM N.A.
RA STRAIN=168 / PB1424;
RC Grand G.,
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR
CC THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN
CC SECONDARY METABOLISM
CC -1- CORFACTOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES
CC (POTENTIAL).
CC -1- SIMILARITY: CONTAINS 5 ACYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcement/>
CC or send an email to license@ebi.ac.uk).

FT DOMAIN 3575 3776 BETA-KETOACYL REDUCTASE 2.
FT DOMAIN 3854 3923 ACYL CARRIER (ACP) 5.
FT DOMAIN 4019 4373 BETA-KETOACYL SYNTHASE 4.
FT BINDING 243 243 PHOSPHOPANTETHEINE (POTENTIAL).
FT BINDING 1723 1723 PHOSPHOPANTETHEINE (POTENTIAL).
FT BINDING 2523 2523 PHOSPHOPANTETHEINE (POTENTIAL).
FT BINDING 2664 2664 PHOSPHOPANTETHEINE (POTENTIAL).
FT BINDING 3886 3886 PHOSPHOPANTETHEINE (POTENTIAL).
SQ SEQUENCE 4427 AA; 493398 MW; 9612515E51AB9F2 CRC64;
Query Match 37.2%; Score 843; DB 1; Length 4427;
Best Local Similarity 41.0%; Pred. No. 7, 2e-50;
Matches 175; Conservative 75; Mismatches 171; Indels 6; Gaps 3;
Qy 1 EPIATGMACRFGVTSADDFWDLSSBDALGSPFTDGMPLDLYDPPRPHGVCYR 60
Db 2763 DAIAIVMSGRIP-GARVREIWDNLVHARNAIRDPIFSH-WVDYKYDVLNKKGVYC 2820
Qy 61 RNCGLPYDAGHPDAEFEGISPREALMDPOQRLLETAMETIEHAGINPTLHGTPGV 120
Db 2821 KSMGMLDIEHPDLPFNIPSEALMDPQHRIFLDEGYAFEDAGNATLNBKKGCV 2880
Qy 121 TGNQDYALRVHNAQSDTDFALVTGTSVSGRISYTFGFGPAVSVDTCSSIVL 180
Db 2881 LGIMSNBYGVMRLRQSRAN---ATGNSFAIAARIPYFLNKGPAIPIDTACSSSIYGT 2936
Qy 181 HLAQALRAGECSMALAGVTWSSPGAFVFERGRGLADGCKA PSAADGTGMEGV 240
Db 2937 HLAQALINKEIDMALVGVGLTLPSTSYMSCEAGMLSPDGCCKAFDNGANFVDEGA 2996
Qy 241 GMLVRLSDAHRNGRVLAVGSAVNODGASNGLTAPGPGQGVIRQALNAGLSAG 300
Db 2997 GALVRLDLDAEDRHHIIGIIGSINDGCKNGITAPSAKQMDLERIYETGHNRE 3056
Qy 301 DVDVAHGTGTTLPPIEQLALATYGGDRAGEGPMGSKVSNVGTGAAAGVAVIK 350
Db 3057 SISYVMHGCTGQGPIDLEALSTVFOETDKQPCAGISVSNIGHTSAAGVAVGVOK 3116
Qy 361 MWALRHGLPRTLHVDSPHYDMSAGAVQLTETVPMPGCGRIARRAGVSFGVSGTN 420
Db 3117 VLLCMHAKTLVPTLNTFTTNEHFEHSPLYVNTLKPWTADGKPRACVSSFGVSGTN 3176
Qy 421 AHVILEE 427
Db 3177 AHVILEE 3183
RESULT 10
MSAS_PENPA STANDARD; PRT; 1774 AA.
ID MSAS_PENPA
AC P22357;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 6-methylsalicylic acid synthase (EC 2.3.1.-) (MSAS).
OS Penicillium patulum (Penicillium griseofulvum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OX NCBI_TaxID=5078;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN=DSM 62862;
RX MEDLINE=9106137; PubMed=2209605;
RA Beck U., Ripka S., Stegner A., Schlitz E., Schweitzer E.;
RT "The multifunctional 6-methylsalicylic acid synthase gene of
RT Penicillium patulum. Its gene structure relative to that of other
RT polyketide synthases." ;
RL Eur. J. Biochem. 192:487-498(1990).
CC -1- FUNCTION: This multifunctional enzyme is a polyketide synthase.
CC It catalyzes a total of 11 steps by seven different component
CC enzymes, in the biosynthesis of the antibiotic patulin.
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + 3 malonyl-CoA + NADPH =
CC 6-methylsalicylic acid + NADP(+) + 3 CoA + 3 CO(2) + H(2)O.

DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR000794; ketoacyl-synt.
 DR InterPro: IPR001601; methyltransf.
 DR InterPro: IPR003880; Pantne atatch.
 DR InterPro: IPR000051; SAM_bind.
 DR Pfam: PF00106; adh_short; 1.
 DR Pfam: PF00109; ketoacyl-synt; 3.
 DR Pfam: PF00550; pp-binding; 4.
 DR Pfam: PF02801; ketoacyl-synt C; 3.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 2.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE: PS50075; ACP_DOMAIN; 4.
 KW Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;
 KM Phosphopantetheine; Multifunctional enzyme; Repeat; Complete proteome.
 FT DOMAIN 295 364 ACYL CARRIER (ACP) 1.
 FT DOMAIN 2190 2258 ACYL CARRIER (ACP) 2.
 FT DOMAIN 2322 2737 BETA-KETOACYL SYNTHASE 2.
 FT DOMAIN 3532 3947 BETA-KETOACYL SYNTHASE 3.
 FT DOMAIN 3410 3483 ACYL CARRIER (ACP) 3.
 FT BINDING 4140 4209 ACYL CARRIER (ACP) 4.
 FT BINDING 327 327 PHOSPHOPANTETHEINE (POTENTIAL).
 FT BINDING 2222 2222 PHOSPHOPANTETHEINE (POTENTIAL).
 FT ACT_SITE 2476 2476 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT BINDING 3446 3446 PHOSPHOPANTETHEINE (POTENTIAL).
 FT ACT_SITE 3690 3690 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT BINDING 4172 4172 PHOSPHOPANTETHEINE (POTENTIAL).
 FT CONFLICT 103 103 E -> V (IN REF. 2).
 FT CONFLICT 276 276 Q -> E (IN REF. 2).
 FT CONFLICT 289 289 T -> S (IN REF. 2).
 SQ SEQUENCE 4273 AA; 477459 MW; 3BBFCFLA250AB5A CRC64;

Query March 35.0%; Score 793; DB 1; Length 4273;
 Best local similarity 39.1%; Pred. No. 1.9e-46;
 Matches 168; Conservative 82; Mismatches 188; Indels 22; Gaps 5;

QY 3 IAIIVGMACRPGVTSADDFWDLISSBODAGCPPTRGMDLTLTYPDPDPCTCYTRN 62
 DB 2322 VAVIGISGRVQART-ABDFNNLKEGNCIEIPKDR-WDKAYVYKKEKESITYTKW 2379
 QY 63 GGFYLDGHPDAEFGISPRALAMDROQLLETANETTHAGINHTL-HGPTGVFT 121
 DB 2380 GGFYLDGHPDAEFGISPRALAMDROQLLETANETTHAGINHTL-HGPTGVFT 121
 QY 122 GTNGDVALRVNMGSTGDPALNGTASVTSRISFEGEPGAVSDTACSSSLVALH 181
 DB 2440 GVMNKNTP-----TGYGWSIANRISTULNFQPSLAVDTACSSSLTALH 2484
 QY 182 LACQALRAGECSMALAGVTVWSSPGAFFVPSRORGLAADGCKAFSAADGTGMEGVG 241
 DB 2485 LALESISGSSDCAIAGSVNLVDPVHYQNLSVNMNLSASDTCKSPDDDDGFPVDEGVG 2544
 QY 242 MLVERLSDAHRNGRVLAIVRGSANVODGASNGLTAPNGPSQQRVTRQALANGLSAGD 301
 DB 2545 AIVLKPRQALADDDHLYGVKAGAINSGCKTNGYTPNPQAQOVKEALERADIPART 2604
 QY 302 VDVAEAGTGTGTDPIEAQALLATYGQDRAGSGRPLWGSVKSVMGHTQAAAGVAVYIM 361
 DB 2605 ISTEAGGTGTDPIEAQALLATYGQDRAGSGRPLWGSVKSVMGHTQAAAGVAVYIM 361
 QY 362 VMALRHGLPRTTLVNDPSRPHVDSAGAVOLLTETVWP---GGEGRLLRAGVSSTGV 417
 DB 2665 LFGKRYQGIAPSLAAGRLNPIERSHNPVYVQQLGEMKRVIGGQVPRRAGLSSTGAG 2724
 QY 418 GTNAHVLLER 427
 DB 2725 GSNHIIILEB 2734

RESULT 12
 PAS_CHICK STANDARD; PRT; 2511 AA.
 ID_PAS_CHICK
 AC P12276;

DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [Includes: EC 2.3.1.38; EC 2.3.1.39;
 DE EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN OR FAS.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archaeoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-11.
 RC STRAIN=White leghorn; TISSUE=Liver;
 RX MEDLINE=95031085; PubMed=7944406;
 RT Huang W.-Y., Chirala S.S., Wakil S.J.;
 RT "Amino-terminal blocking group and sequence of the animal fatty acid
 RT synthase.";
 RT Arch. Biochem. Biophys. 314:45-49(1994).
 RN [2]
 RP SEQUENCE OF 75-1775 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89282777; PubMed=2734291;
 RA Holzer K.P., Liu W., Hammes G.G.;
 RT "Molecular cloning and sequencing of chicken liver fatty acid
 RT synthase cDNA.";
 RT Proc. Natl. Acad. Sci. U.S.A. 86:4387-4391(1989).
 RN [3]
 RP SEQUENCE OF 1568-2512 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=8919426; PubMed=2917973;
 RA Chirala S.S., Kasturi R., Pazirandeh M., Stojlov D.T., Huang W.-Y.,
 RA Wakil S.J.;
 RA "A novel cDNA extension procedure. Isolation of chicken fatty acid
 RT synthase cDNA clones.";
 RT J. Biol. Chem. 264:3750-3757(1989).
 RN [4]
 RP SEQUENCE OF 1752-2512 FROM N.A.
 RX MEDLINE=88320436; PubMed=2842766;
 RA Yuan Z., Liu W., Hammes G.G.;
 RT "Molecular cloning and sequencing of DNA complementary to chicken
 RT liver fatty acid synthase mRNA.";
 RT Proc. Natl. Acad. Sci. U.S.A. 85:6328-6331(1988).
 RN [5]
 RP SEQUENCE OF 2202-2512 FROM N.A.
 RX MEDLINE=89088152; PubMed=3207710;
 RA Kasturi R., Chirala S.S., Pazirandeh M., Wakil S.J.;
 RT "Characterization of a genomic and cDNA clone coding for the
 RT thioesterase domain and 3' noncoding region of the chicken liver
 RT fatty acid synthase gene.";
 RT Biochemistry 27:7778-7785(1988).
 RN [6]
 RP SEQUENCE OF 2121-2209.
 RX MEDLINE=89192401; PubMed=2648999;
 RA Huang W.-Y., Scoop J.K., Wakil S.J.;
 RT "Complete amino acid sequence of chicken liver acyl carrier protein
 RT derived from the fatty acid synthase.";
 RT Arch. Biochem. Biophys. 270:92-98(1989).
 RN [7]
 RP SEQUENCE OF 2209-2508.
 RC STRAIN=White leghorn;
 RX MEDLINE=89088151; PubMed=3207709;
 RA Yang C.-Y., Huang W.-Y., Chirala S.S., Wakil S.J.;
 RT "Complete amino acid sequence of the thioesterase domain of chicken
 RT liver fatty acid synthase.";
 RT Biochemistry 27:7773-7777(1988).
 RN [8]
 RP SEQUENCE OF 667-674 AND 1698-1709.
 RX MEDLINE=89323081; PubMed=2751995;
 RA Chang S.I., Hammes G.G.;
 RT "Amino acid sequences of pyridoxal 5'-phosphate binding sites and
 RT fluorescence resonance energy transfer in chicken liver fatty acid
 RT synthase.";
 RT Biochemistry 28:3781-3788(1989).

CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl]-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxypalmitoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + NADPH.
 CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC -1- CAUTION: REF.3 AND REF.5 SEQUENCES DIFFER FROM THAT SHOWN FROM
 CC POSITION 2351 ONWARDS DUE TO A FRAMESHIFT.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J04485; ABA46389.1; -
 DR EMBL: J03860; AAA48767.1; -
 DR EMBL: J02839; AAA82106.1; ALT_SEQ.
 DR PIR: A33918; XYCHA.
 DR PIR: A32015; A32015.
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR002085; Adh_zn family.
 DR InterPro: IPR007994; Ketoacyl-synt.
 DR InterPro: IPR003880; Prantne_attach.
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00109; ketoacyl-synt; 1.
 DR Pfam: PF00550; pp-binding; 1.
 DR Pfam: PF00698; Acyl_transf; 1.
 DR Pfam: PF00975; Thioesterase; 1.
 DR Pfam: PF02801; ketoacyl-synt_C; 1.
 DR PROSITE: PS00012; PHOSPHOPANTHEINE, 1.
 DR PROSITE: PS00606; B KETOACYL SYNTHASE; 1.
 DR PROSITE: PS50075; ACP DOMAIN; 1.
 KW Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 KW Transferase; Hydrolase; Oxidoreductase; Lyase; NADP; Acetylation;
 KW Alternative splicing; Pyridoxal phosphate.
 KM
 FT INIT MET 0
 FT DOMAIN 1 2411 BETA-KETOACYL SYNTHASE.
 FT DOMAIN 427 815 ACYL AND MALONYL TRANSFERASES.
 FT DOMAIN 1638 1866 ENOYL REDUCTASE.
 FT DOMAIN 1867 2119 BETA-KETOACYL REDUCTASE.
 FT DOMAIN 2124 2180 ACYL CARRIER (ACP).
 FT DOMAIN 2209 2511 THIOESTERASE.
 FT MOD RES 1 1 ACETYLTATION.
 FT ACT_SITE 160 160 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT_SITE 579 579 ACYL/MALONYL TRANSFERASES (BY
 FT ACT_SITE 877 877 SIMILARITY).
 FT ACT_SITE 1674 1691 BETA-HYDROXYACYL DEHYDRATASE (BY
 FT BINDING 1707 1707 SIMILARITY).
 FT NP_BIND 1888 1903 NADP (R).
 FT BINDING 2157 2157 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT ACT_SITE 2308 2308 NADP (R).
 FT BINDING 1903 1903 THIOESTERASE (BY SIMILARITY).
 FT BINDING 2157 2157 THIOESTERASE (BY SIMILARITY).
 FT ACT_SITE 2308 2308 THIOESTERASE (BY SIMILARITY).

FT ACT_SITE 2481 2481 THIOESTERASE (BY SIMILARITY).
 FT VARSPIC 2348 2348 T -> TCCFSFLP (IN ISOFORM 1).
 FT CONFLICT 77 78 OL -> PV (IN REF. 2).
 FT CONFLICT 116 116 L -> A (IN REF. 2).
 FT CONFLICT 675 675 R -> S (IN REF. 2).
 FT CONFLICT 1169 1169 K -> N (IN REF. 2).
 FT CONFLICT 1178 1178 A -> T (IN REF. 2).
 FT CONFLICT 1191 1191 R -> H (IN REF. 2).
 FT CONFLICT 1198 1198 P -> L (IN REF. 2).
 FT CONFLICT 1286 1287 DN -> ND (IN REF. 2).
 FT CONFLICT 1372 1372 K -> E (IN REF. 2).
 FT CONFLICT 1533 1533 C -> Y (IN REF. 2).
 FT CONFLICT 1577 1577 W -> R (IN REF. 2).
 FT CONFLICT 1685 1696 QAAIATLSMGC -> ASSHCHRLHGLA
 FT CONFLICT 1732 1732 (IN REF. 2; AAA48767).
 FT CONFLICT 1745 1745 O -> E (IN REF. 2).
 FT CONFLICT 2511 2511 S -> N (IN REF. 2).
 SQ SEQUENCE 2511 AA; 274648 MW; 622039DAC8315D3F CRC64;
 Query Match 31.5%; Score 713; DB 1; Length 2511;
 Best Local Similarity 36.9%; Pred. No. 3,3e-41;
 Matches 158; Conservative 83; Mismatches 159; Indels 28; Gaps 9;
 QY 1 EPIAIVGMAQRFPGVTSADDFWDLISSEODATIGFPTDGMDDLTYDPDPHPTC-Y 59
 DB 1 EDVVIAGIADKLPESS-ENLQEFWENLNGVDWV--TEDDRW-----KPGIYGL 46
 QY 60 TRNGFLYDGHDPDAERFGISPREALMDPQOQLLETARETEHAGINPHTHAGPTGV 119
 DB 47 PKRNGKIKDKTKKFDASFGVHPQAHMDQLLEVSFEALDGLGINTPALRGDTGV 106
 QY 120 FTGNGQDYALRYVNAQSGTDGFPALNGTAGVISGHSISYFGEPAVSVDATCSSSLVA 179
 DB 107 WVGASGEALFALSDPBEILGYSMTCQAMLANISYVDYDTGSLTIDTCSSSLMA 166
 QY 180 LHLAQLALRAGECSSMALAGVTWSSPFAVEFSRQGLAADCHCAFSAADGTWKEG 229
 DB 167 LENAAYKAIIRHQCSAALVGVNITLKPNITSVQPMKLGMLSPDACRAFDVSGNGYCRSEA 226
 QY 240 VGNLIVERRSDARHNGRYVAVVRGSAVNDG-AVNGLTLPNPSOORVRLQALNAGLS 298
 DB 227 VVVVLTKKSWA---KRVATTVNAGSNTDGKEQGVTFPSGEMQOQLVSLYREGIK 282
 QY 299 AGDVDAVEAHGTGTGDTPIEAQALLATYQODRAGEGPTLGSVKSNGHTQAAGVAGV 358
 DB 283 PGDVEYEAHGTGTGDTGDPGVNGVNVFCQ--CEREPLITGSKSMGHPSPASGLAAL 340
 QY 359 IKRWMAIRHGLPRTTHVDEPSPHVD-WSAGAVQLLTETVVPWPGEGRLRAAGVSFGVS 417
 DB 341 AKYILSLHGLMAPNLHFNPNPDIPALHDGSLKVKCKPVPVKG-----LVINSRFG 395
 QY 418 GTNAHYTL 425
 DB 396 GSNNAHYTL 403
 RESULT 13
 STCA_EMENT STANDARD; PRT; 2181 AA.
 ID STCA_EMENT
 AC Q12397;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative sterigmatacystin biosynthesis polyketide synthase (PKS).
 GN Emericella nidulans (Aspergillus nidulans).
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FGSC 26;
 RX MEDLINE=96202293; PubMed=8643646;

RA Brown D.W., Yu J.-H., Kellar H.S., Fernandes M., Nesbitt T.C.,
 RA Kellier N.P., Adams T.H., Leonard T.J.,
 RT "Twenty-five coregulated transcripts define a sterigmatocystin gene
 RT cluster in *Aspergillus nidulans*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:1418-1422(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FGSC 4;
 RX MEDLINE=95370159; PubMed=7642507;
 RA Yu J.-H., Leonard T.J.,
 RT "Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a
 RT novel type 1 polyketide synthase.";
 RL J. Bacteriol. 177:4792-4800(1995).
 CC -1- FUNCTION: INVOLVED IN THE SYNTHESIS OF THE POLYKETIDE NUCLEUS OF
 CC STERIGMATOCYSTIN FROM HEXANOYL-COA AND SEVEN MALONATES.
 CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETHEINES
 CC (POTENTIAL).
 CC -1- PATHWAY: Sterigmatocystin biosynthesis; first step.
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1eb-sib.ch/announce/>
 CC or send an email to license@1eb-sib.ch).
 CC -----
 CC EMBL: U34740; AAC49191.1; -;
 DR EMBL: L39121; AAA81586.1; -;
 DR HSSP: P25715; IMLA.
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantane attach.
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00109; ketoacyl-synt; 1.
 DR Pfam: PF00550; pp-binding; 1.
 DR Pfam: PF00698; Acyl transfer; 2.
 DR Pfam: PF00975; Thioesterase; 1.
 DR Pfam: PF02801; ketoacyl-synt; C; 1.
 DR PROSITE: PS50075; ACP_DOMAIN; 2.
 KW Transferase; Acyltransferase; Phosphopantetheine; Repeat;
 KW Multifunctional enzyme.
 FT DOMAIN 383 814 BETA-KETOACYL SYNTHASE.
 FT DOMAIN 884 1209 ACYL/MALONYL TRANSFERASES.
 FT DOMAIN 1706 1777 ACYL CARRIER (ACP) 1.
 FT DOMAIN 1830 1901 ACYL CARRIER (ACP) 2.
 FT DOMAIN ? 2181 THIOESTERASE.
 FT ACT_SITE 552 552 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT_SITE 978 978 ACYL/MALONYL TRANSFERASES (BY
 FT SIMILARITY).
 FT BINDING 1738 1738 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT BINDING 1862 1862 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT ACT_SITE 2028 2028 THIOESTERASE (BY SIMILARITY).
 SQ SEQUENCE 2181 AA; 238831 MW; 5A3E5712AA9AD942 CRC64;
 Query March 31.0%; Score 702.5; DB 1; Length 2181;
 Best Local Similarity 37.0%; Pred. NO. 1.5e-40;
 Matches 162; Conservative 74; Mismatches 185; Indels 17; Gaps 8;
 QY 3 IAIYGMACRFPGVTSADDFWDLISSECDALIGFPTRGMDLDTLYDPDPDHPGTCYTRN 62
 DB 383 LAIVSMGRFP-EAPSTDSFMDLLYKGLDVCKEVL-RRMDVKTHVDPSCGARRKXKTRW 440
 QY 63 GGFAYDAGHDAEFGISPREALAMPQORLLLETAMETIEHAGINPTLHGIP---TGV 119
 DB 441 GCWLDPAGBEPDPFRFSPISPEKAPMDPAORMALMSTEAEMERGIVDTPTSTORRIGV 500
 QY 120 FTGNNGDYALRVNNAAGSTGSPALITGPAAGSVISGRISYGPGEFPAVSVDATCASSIVA 179
 DB 501 FHGTSTSDW--METNTAKONIDTYITFGNNGFIFGRINFCFEFGSPYSNDTACSSSLAA 558
 QY 180 LHLAQUALRAGECSMALAGVTWSSPGAFFVEFSRORGLAADGHCKAFSAADGTGWGEG 239

DB 559 IHLACNSLWKGDDDTAVAGGTNMIFFPDGHTGLDKGFILSRNTCKAFDPAADGCRABEG 618
 QY 240 VGMALVERLSDAHRNHRVLAVRGSAVNDGASNLTLTPNGSQGRVIRQALANGLSA 239
 DB 619 VGVTFIKRLLEDALAENDPILATTLIDIKTNNSAMSDMTFRPKAQIDNNSSALLSTAGISP 678
 QY 300 GDVDAVEAGTGTTLGDPIDPAQALATYGGD---RAGEGFLMGSVKSVNGHTQAAAGVA 356
 DB 679 LDI SYEMHGTGTQVGDVAVEMESVLSFAPDEFRRDRKLYGSAKANGHEGVSQVT 738
 QY 357 GVIRKVMALRHGILLPTLHVDESPHV----DMSAGAVQLLTETVPMGSGRLRRAGV 411
 DB 739 SLIKVILMMKNDTILPHCGI-KPGRINRNVPLPARNVHIAEPKWPRTD-TPRRVLI 796
 QY 412 SSGVSGTANAVLLEAP 429
 DB 797 NNPSAAGNTAVLVEDAP 814
 RESULT 14
 PKSI_ASPPA STANDARD; PRT; 2109 AA.
 ID PKSI_ASPPA
 AC Q12053;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aflatoxin biosynthesis polyketide synthase (PKS).
 GN PKS1.
 OS *Aspergillus parasiticus*.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; *Aspergillus*.
 OX NCBI_TaxID=5067;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL 2999;
 RX MEDLINE=96042102; PubMed=7592391;
 RA Peng G.H., Leonard T.J.,
 RT "Characterization of the polyketide synthase gene (pksl) required
 RT for aflatoxin biosynthesis in *Aspergillus parasiticus*.";
 RL J. Bacteriol. 177:6246-6254(1995).
 CC -1- FUNCTION: INVOLVED IN THE SYNTHESIS OF THE POLYKETIDE NUCLEUS OF
 CC AFLATOXIN FROM HEXANOYL COA AND SEVEN MALONATES.
 CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETHEINE (By
 CC similarity).
 CC -1- PATHWAY: Aflatoxin biosynthesis; first step.
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1eb-sib.ch/announce/>
 CC or send an email to license@1eb-sib.ch).
 CC -----
 CC EMBL: L42765; AAC41675.1; -;
 DR EMBL: L42765; AAC41674.1; -;
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantane attach.
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00109; ketoacyl-synt; 1.
 DR Pfam: PF00550; pp-binding; 1.
 DR Pfam: PF00698; Acyl transfer; 1.
 DR Pfam: PF00975; Thioesterase; 1.
 DR Pfam: PF02801; ketoacyl-synt; C; 1.
 DR PROSITE: PS50075; ACP_DOMAIN; 1.
 KW Transferase; Acyltransferase; Phosphopantetheine;
 KW Multifunctional enzyme.
 FT DOMAIN ? 805 BETA-KETOACYL SYNTHASE.
 FT DOMAIN 374 ? ACYL/MALONYL TRANSFERASES.
 FT DOMAIN 1714 1785 ACYL CARRIER (ACP).

FT DOMAIN ? 2109 THIOESTERASE.
 FT ACT_SITE 543 543 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT_SITE 993 993 ACYL/MALONYL TRANSFERASE (BY
 SIMILARITY).
 FT BINDING 1746 1746 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT ACT_SITE 1937 1937 THIOESTERASE (BY SIMILARITY).
 SQ SEQUENCE 2109 AA; 230715 MW; CB701372A16D8551 CRC64;
 Query Match 30.2%; Score 684.5; DB 1; Length 2109;
 Best Local Similarity 35.4%; Pred. No. 2.5e-39;
 Matches 158; Conservative 74; Mismatches 181; Indels 33; Gaps 8;
 QY 3 IAIWMACTPFGVTSADDFWDLISBQDAIGCFPTDGMWDLTLTYPDPDHPCTCYTRN 62
 DB LAIVSMGSRFPESPFT-ESFWDLYKGLDVCKEVPR-RMDIINHVPDGSARKKATKM 431
 QY 63 GGFYDAGHPAEFFGSGSPREALMDPQRLLTETAMETTHAGINPTLHGTGTP--TGV 119
 DB GCWLDGSGDPPREFGIGSPKEAPQMDPAORMALMSTYEMERAGIIVDPPTSTORDRIGV 491
 QY 120 PTGTNGDVALRVHNAAGSTGDPALVTGAGSVISGRISTYTFGFGPAVSVTACSSSLVA 179
 DB 492 FHGVTSDMW--MERTNAQNTITTYITGNGRGPITGRINFCPEFGPSTYNDTACSSSLA 549
 QY 180 LHACQALRAGECSMALAGVTVNWSPPGAFVEFSRQGLAADGCKAFSAADGTGWEGB 239
 DB 550 IHLACNSLMRGDCDTAVAGGTMNMTYTPDGHGTGLKGFPLSRGCKPDKADGYCBAEG 609
 QY 240 VGMILVERLSPAHNGHRVLAIVGSAVNOGASNGULTAPNGPQQRVIRALANAGISA 299
 DB 610 VGVTFIRKLEBDALNDPILGVIIDAKTNHSMESMTFRPVGAQIDMTALMTTGLHP 669
 QY 300 GDVDAVEAHGCTTGLGPPIEAQALATYQD--RAGEGPTLGSVSNVGHTOAAAGVA 356
 DB 670 NDESYIEHGGTGVGDAVEMESULSFAPSETRKADQPLFVSAAKAVHGEGVSVT 729
 QY 357 GVIRKMAVRHGLPRTLHVPDESPHV-----DMSAGAVQLTETVPMRGE 403
 DB 730 SLIKVLMWQHDITP-----PHGIRKSGKINNFPDLGARVNHIAEPMPWRTH 780
 QY 404 GLRLRAGVSSFGVSGTNAHVLEEAR 429
 DB 781 -TPRRVLINNFSAAGNTALIVEDAP 805
 RESULT 15
 ID_FAS_RAT STANDARD; PRT; 2505 AA.
 AC P12785; Q64717; O09187; O09190;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [includes: EC 2.3.1.36; EC 2.3.1.39;
 EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=89240686; PubMed=2717611;
 RA Amy C.M., Witkowski A., Naggett J., Williams B., Randhava Z.,
 RA Smith S.;
 RT "Molecular cloning and sequencing of cDNAs encoding the entire rat
 RT fatty acid synthase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3114-3118(1989).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=93075999; PubMed=1339331;
 RA Beck K.F., Schlegelmann R., Stathopoulos I., Klein H., Hoch J.,
 RA Schweitzer M.;

RT "The fatty acid synthase (FAS) gene and its promoter in Rattus
 RT norvegicus.";
 RT DNA Seq. 2:359-386(1992).
 RL [3]
 RP SEQUENCE OF 75-2505 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Mammary gland;
 RX MEDLINE=89128431; PubMed=2915923;
 RA Schweitzer M., Takabayashi K., Beck K.F., Schlegelmann R.;
 RT "Rat mammary gland fatty acid synthase: localization of the
 RT constituent domains and two functional polyadenylation/termination
 RT signals in the cDNA.";
 RL Nucleic Acids Res. 17:567-586(1989).
 RL [4]
 RP SEQUENCE OF 2085-2505 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=89087240; PubMed=2891707;
 RA Naggett J., Witkowski A., Mikkelsen J., Smith S.;
 RT "Molecular cloning and sequencing of a cDNA encoding the thioesterase
 RT domain of the rat fatty acid synthetase.";
 RL J. Biol. Chem. 263:1146-1150(1988).
 RL [5]
 RP SEQUENCE OF 1921-2324 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=8724646; PubMed=3109907;
 RA Witkowski A., Naggett J., Mikkelsen J., Smith S.;
 RT "Molecular cloning and sequencing of a cDNA encoding the acyl carrier
 RT protein and its flanking domains in the mammalian fatty acid
 RT synthetase.";
 RL Eur. J. Biochem. 165:601-606(1987).
 RL [6]
 RP FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 RP LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 RP THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 RP ACYL CARRIER PROTEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxypalmitoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-
 CC carrier protein] + oleate.
 CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M76767; AAA57219.1; -
 CC EMBL; X62888; CAA44679.1; -
 CC EMBL; X62889; CAA44680.1; -
 CC EMBL; X13415; CAA31780.1; -
 CC EMBL; X13527; CAA31882.1; -
 CC EMBL; J03514; AAA41144.1; -
 CC PIR; A30313; XYRTFA.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh zn family.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Ppancne attach.
 DR InterPro; IPR000051; SAM bind.
 DR InterPro; IPR001031; Thioesterase.

Job time : 5.68929 secs.

DR pfam; PF00107; adh_zinc; 1.
DR pfam; PF00109; ketoacyl-lyase; 1.
DR pfam; PF00550; pp-binding; 1.
DR pfam; PF00698; acyl transferase; 1.
DR pfam; PF00975; thioesterase; 1.
DR pfam; PF02801; ketoacyl-lyase; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
DR PROSITE; PS00606; B KETOACYL SYNTHASE; 1.
DR PROSITE; PS50075; ACP DOMAIN; 1.
KM Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
KW Hydroxylase; Oxidoreductase; Transferase; Lyase; NADP;
KW Pyridoxal phosphate; 413
FT DOMAIN 1
FT DOMAIN 429 817 BETA-KETOACYL SYNTHASE.
FT DOMAIN 1629 1857 ACYL AND MALONYL TRANSFERASES.
FT DOMAIN 1858 2113 ENOYL REDUCTASE.
FT DOMAIN 2118 2174 BETA-KETOACYL REDUCTASE.
FT DOMAIN 2202 2505 ACYL CARRIER (ACP).
FT ACT_SITE 161 161 THIOESTERASE.
FT ACT_SITE 581 581 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT NP_BIND 1665 1682 MALONYLTRANSFERASE (BY SIMILARITY).
FT BINDING 1698 1698 NADP (ER).
FT NP_BIND 1765 1780 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT BINDING 2151 2151 NADP (KR).
FT ACT_SITE 2302 2302 PHOSPHOPANTHETHEINE (BY SIMILARITY).
FT ACT_SITE 2475 2475 THIOESTERASE (BY SIMILARITY).
FT ACT_SITE 2475 2475 THIOESTERASE (BY SIMILARITY).
FT ACT_SITE 878 878 BETA-HYDROXYACYL DEHYDRATASE (BY SIMILARITY).
FT ACT_SITE 871 871 S -> P (IN REF. 3).
FT CONFLICT 1967 1968 MV -> IL (IN REF. 5).
FT CONFLICT 2085 2085 C -> P (IN REF. 4).
FT CONFLICT 2106 2106 A -> V (IN REF. 1 AND 5).
FT CONFLICT 2296 2296 Y -> H (IN REF. 1 AND 5).
SQ SEQUENCE 2505 AA; 272647 MW; 5810EC13D37F3114 CRC64;

Query Match 30.1%; Score 682; DB 1; Length 2505;
Best Local Similarity 34.9%; Pred. No. 4.5e-39;
Matches 153; Conservative 92; Mismatches 159; Indels 34; Gaps 10;

QY 1 EPIAIVGMACRFPGCVTSADDFMDLISSEODATCGFPTDRGMDLDTLYDPDPHRTCY- 59
DB 2 EEVVIAGMSGKLPES-ENLOEFW-----ANLIGG-----VDWYTDDDRKKAGLYG 46
QY 60 -TRNGCFLYAGHGFDAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLHGPTG 118
DB 47 LPKRSGLKDLKSKDASFFGHPKQAHMTBQRLLEVESEALVDGGINPASLRGNTG 106
QY 119 VFTGTNGQDYALRVHNAQSTGDFALTGTAISVIGRIISYTFEGEPASVDTACSSSLV 178
DB 107 VMVGVSSEASSEALSROPELTILGYSVMGQORAMMANRLSPFFDFKPSIALDTACSSSL 166
QY 179 ALHLACQALRAGCESMALAGSVTMSFGAFVEFSRQGLAAGHCKAFSAADGTGME 238
DB 167 ALQNAVQAIRSGECPAIVGGINLLKPNISVQFMKGLSPDGTCSRFPDSSGNGYCRAE 226
QY 239 GVGMLVERLSDAHRNGRVLAVVRGSAVNODGA-SNGLTAPNGPSOORVIRQALANAGL 297
DB 227 AVAVAVLTKKSLSA-----RRYATITLNGNTIDCKEQGVTFPSEAOEQLIRSLYQPGV 282
QY 298 SAGVDVAEAGHTGTTLGDPLEAQAALATYQODRAGEPLMISVKSNGHTQAAGVAG 357
DB 283 APESLEIYEIAHGTGTGVGDPELNGITRSLCAFR--QSPILLIGSTKSNMGHPSPASGLAA 340
QY 358 VIKVMMLRRLGLPRLTHVDESPHYD-WSAGAVQLLTETVPMWPGSGRLRRAGVSSFGV 416
DB 341 LTKVLLSLKLVGVAPNHLFHPNPDEIPALLDGRLOVVDRLPYRGG-----YVGINSGF 395
QY 417 SGTAAVYL-----EADA 430
DB 396 GGANVHYILOPNTQQA 413

Search completed: June 17, 2003, 13:02:37

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 17, 2003, 12:55:32 ; Search time 9.60187 Seconds
(without alignments)
4305.183 Million cell updates/sec

Title: US-09-914-286-3_COPY_2086_2515
Perfect score: 2265
Sequence: 1 EPIAIVGMACRPFQGVTSAD.....VSSFGVSGTNAHVLEBAPA 430

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1654	73.0	3519	2	polyketide synthase
2	1638	72.3	2723	2	probable polyketid
3	1636	72.2	5069	2	riofamycin polyketi
4	1628.5	71.9	3739	2	polyketide synthase
5	1622.5	71.6	4613	2	polyketide synthase
6	1611	71.1	4735	2	riofamycin polyketi
7	1591.5	70.3	3413	2	riofamycin polyketi
8	1590	70.2	7576	2	FK506 polyketide
9	1586	70.0	10223	2	polyketide synthase
10	1581.5	69.8	8563	2	polyketide synthase
11	1574	69.5	2100	2	probable polyketid
12	1570.5	69.3	1728	2	riofamycin polyketi
13	1567.5	69.2	1763	2	riofamycin polyketi
14	1563	69.0	1937	2	probable polyketid
15	1561.5	68.9	1562	2	polyketide synthase
16	1559	68.8	6260	2	polyketide synthase
17	1558.5	68.8	4151	2	probable polyketid
18	1556	68.7	2103	2	polyketide synthase
19	1554	68.6	6420	2	polyketide synthase
20	1550	68.4	496	2	probable polyketid
21	1544.5	68.2	3573	2	erythronolide synt
22	1538	67.9	1346	2	polyketide synthase
23	1533.5	67.7	1762	2	probable polyketid
24	1521	67.2	1602	2	probable polyketid
25	1497.5	66.1	2126	2	probable polyketid
26	1495.5	66.0	3491	2	probable 6-deoxyer
27	1425.5	62.9	3172	2	erythronolide synt
28	1425.5	62.9	3178	2	6-deoxyerythronol
29	1170.5	51.7	2124	2	polyketide synthase

30	1164	51.4	2201	2	S73014	polyketide synthase
31	1163.5	51.4	2518	2	A12140	polyketide synthase
32	1151.5	50.8	2188	2	A70984	probable polyketid
33	1148	50.7	1587	2	AB2012	hypothetical prote
34	1139.5	50.3	1827	2	B70984	probable polyketid
35	1135.5	50.1	1815	2	S73021	polyketide synthase
36	1135.5	50.1	1822	2	F87203	polyketide synthase
37	1110	49.0	1446	2	S73013	polyketide synthase
38	1110	49.0	1540	2	H87203	polyketide synthase
39	1109.5	48.8	1538	2	E70874	probable mbcC prot
40	1104.5	48.6	1293	2	T30871	probable ppsB prot
41	1101	48.3	2478	2	AH2140	orellanic acid sy
42	1095	47.9	1774	2	T17421	polyketide synthase
43	1085.5	47.1	1876	2	C70749	probable ppsA prot
44	1066	46.5	1871	2	A87204	polyketide synthase
45	1054	46.5	1871	2	A87204	polyketide synthase

ALIGNMENTS

RESULT 1

S43048
polyketide synthase type I - Streptomyces antibioticus
N/Contains: acyl carrier protein; acyltransferase; ketoreductase; ch1
C/Species: Streptomyces antibioticus
C/Date: 13-Jan-1995 #sequence_revision 06-Dec-1996 #text_change 26-May-2000
C/Accession: S43048, S41729
R/Swan, D.G.; Rodriguez, A.M.; Vilches, C.; Mendez, C.; Salas, J.A.
Submitted to the EMBL Data Library, February 1993
A/Reference number: S43048
A/Accession: S43048
A/Molecule type: DNA

A/Residues: 13519 <SMA>
A/Cross-references: EMBL: L09654; NID: G153407; PID: AAA19695.1; PID: G153408
R/Swan, D.G.; Rodriguez, A.M.; Vilches, C.; Mendez, C.; Salas, J.A.
Mol. Gen. Genet. 242, 358-362, 1994
A/Title: Characterisation of a Streptomyces antibioticus gene encoding a type I polyket
A/Reference number: S41729; PMID: 94150470; PMID: 8107683
A/Accession: S41729
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1683-3238, 3273-3303, 'K', 3305-3407, 'T', 3409-3462, 'Y', 3464-3516, 'E', 3518-3519
A/Cross-references: EMBL: L09654

C/Genetics:
A/Start codon: GTG
C/Superfamily: [acyl-carrier-protein] S-malonyltransferase homology; 3-oxoacyl-[acyl-ca
ogy; short-chain alcohol dehydrogenase homology
C/Keywords: antibiotic biosynthesis; carrier protein
F:59-462/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F:102-1381/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F:1489-1560/Domain: acyl carrier protein homology <ACP1>
F:1708-2111/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F:2221-2502/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F:2858-3037/Domain: short-chain alcohol dehydrogenase homology <SAD2>
F:3143-3214/Domain: acyl carrier protein homology <ACR2>
F:3305-3500/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>

Query Match 73.0%; Score 1654; DB 2; Length 3519;
Best Local Similarity 73.3%; Pred. No. 5.8e-106;
Matches 315; Conservative 40; Mismatches 75; Indels 0; Gaps 0;

QY	1	EPIAIVGMACRPFQGVTSADDFDILISSBODAGFPPTDRCMDLTYDDPPHPTGCTT	60
DB	1687	EALAITAMSCRFQGDIDSPEDLWFLAEGDAVAGLPBDEGMIDLALYHDDPDPPTTYV	1746
QY	61	RNGGFLYDAGHPAEPFGISPREALAMPDQORLLTFAMETTHAGINPTLTGTPGVF	120
DB	1747	REBAFRYDAQFAGFGISPREALAMPDQORLLTFAMETTHAGINPTLTGTPGVF	1806
QY	121	TGTNGDYALRVNAGOSTDGFALTGTAGSVISGRISYTFGFGPAVSVDTACSSSLVAL	180

Db 1807 IGAGHGGYDPPKRAPESSVAGYLLTGTASAVLISGRISYTFGLIEGPAVTVDTACSSSLVAL 1866
Qy 181 HLACQALRAGCSMALAGVTVMSSPGAFAVEFSRQGLAADGHKAPSAADGTGWEQGV 240
Db 1867 HLAVALRGRBCSLAIGQVAVMSTPDHAFVFSRQGLADGRCAFAAADGKGWGGV 1926
Qy 241 GMLIVERLSDAHRNGHRLAVVRSAAVQDASNGLTAPNGSQORVIRQALANAGLSAG 300
Db 1927 SLILLERLSDAHRNGHRLAVVRSAAVQDASNGLTAPNGSQORVIRQALADAGLAPA 1986
Qy 301 DVDAVEAHGTTGTDPIEAQALLATYQDRAEGPLMLGSVKSNGVHTQAAAGVAVIK 360
Db 1987 DVDAVEAHGTTGTDPIEAQALLATYQDRAEGPVMVLSGVKSNGVHTQAAAGVAVIK 2046
Qy 361 MVMALRHGLPRTLHVDEPSPHVMSAGAVOLLTETVPMGEGRLRAGVSSFGVSGTN 420
Db 2047 MVLALGRGVKTLHVDEPSPHVMSAGAVELLTERPMEPRLRAGISARGVSGTN 2106
Qy 421 AHVILEEAP 430
Db 2107 AHVILEEAP 2116

RESULT 2

probable polyketide synthase module 1 - Streptomyces hygroscopicus
C/Species: Streptomyces hygroscopicus
C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 17-Nov-2000
C/Accession: T03221
R/Ruan, X.; Steasi, D.; Lax, S.; Katz, L.
Gene 203, 1-9, 1997
A/Title: A second type-I PKS gene cluster isolated from Streptomyces hygroscopicus ATCC
A/Reference number: Z14848; MUID:98085969; PMID:9426000
A/Accession: T03221
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2723 <RUA>
A/Cross-references: EMBL:AF007101; NID:g2624946; PIDN:AAC38061.1; PID:g2624948
A/Experimental source: ATCC 29253
C/Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase
C/Keywords: carrier protein
F/34-504/Domain: acetate-CoA ligase homology <ACL>
F/81-949/Domain: acyl carrier protein homology <ACP1>
F/996-1397/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F/1509-1794/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
F/2568-2633/Domain: acyl carrier protein homology <ACP2>

Query Match 72.3%; Score 1638; DB 2; Length 2723;

Best Local Similarity 71.3%; Pred. No. 5.4e-105;

Matches 306; Conservative 49; Mismatches 72; Indels 2; Gaps 1;

Qy 1 EPIAIVMACRFPFGVTSADDFMDLISSEODALIGFPTDGMWDLDTLYDPDPHPTCYT 60
Db 975 EPIAIVMACRLPGVTSPEELWELSSGGISGFPTDGMWDLHLHPDPHPTCYT 1034
Qy 61 RNGGFLYDAGHPDAEPFGISPREALAMPPOORLLLETAMETIHAAGINPTLHGTPTGVF 120
Db 1035 REGGFLHDAGDFDALFGISPREALAMPPOORLLLETAMETIHAAGINPTLHGTPTGVF 1094
Qy 121 TGNNGDYALRVHNAAGSTDFGALTGTAGSVISGRISYTFGREGPAVSVDTACSSSLVAL 180
Db 1095 AGVAVHYIYSLRITVPELBEGLSSGSGSVSGVSYTFGREGPAVSVDTACSSSLVAL 1154
Qy 181 HLAQALRAGCSMALAGVTVMSSPGAFAVEFSRQGLAADGHKAPSAADGTGWEQGV 240
Db 1155 HLAVALRGRBCSLAIGQVAVMSTPDHAFVFSRQGLADGRCAFAAADGKGWGGV 1214
Qy 241 GMLIVERLSDAHRNGHRLAVVRSAAVQDASNGLTAPNGSQORVIRQALANAGLSAG 300
Db 1215 ATLIVERLSDAHRNGHRLAVVRSAAVQDASNGLTAPNGSQORVIRQALANAGLSAG 1274
Qy 301 DVDAVEAHGTTGTDPIEAQALLATYQDRAEGPLMLGSVKSNGVHTQAAAGVAVIK 360

Db 1275 DIDLEAHGTTGTDPIEAQAVLAVMYGQDR--DLFVWLSGLKSNIGHTQAAAGVAVIK 1332
Qy 361 MVMALRHGLPRTLHVDEPSPHVMSAGAVOLLTETVPMGEGRLRAGVSSFGVSGTN 420
Db 1333 AVLSMRGVKPSLHDEPPEVDMASGIVELLAEARDMFGADBRPRRAGVSSFGVSGTN 1392
Qy 421 AHVILEEAP 429
Db 1393 AHVILEEAP 1401

RESULT 3

rifamycin polyketide synthase modules 4-6 - Amycolatopsis mediterranei
C/Species: Amycolatopsis mediterranei
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
C/Accession: T17464
R/Schupp, T.
submitted to the EMBL Data Library, December 1997
A/Reference number: 218802
A/Accession: T17464
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-5069 <SCH>
A/Cross-references: EMBL:AF223012; NID:e1227119; PID:e1227121; PIDN:CAA11036.1
A/Experimental source: strain LBG A3136
C/Superfamily: acyl carrier protein homology
C/Keywords: carrier protein
F/1631-1702/Domain: acyl carrier protein homology <ACP1>
F/3238-3309/Domain: acyl carrier protein homology <ACP2>
F/4939-5010/Domain: acyl carrier protein homology <ACP3>

Query Match 72.2%; Score 1636; DB 2; Length 5069;

Best Local Similarity 72.4%; Pred. No. 1.6e-104;

Matches 315; Conservative 43; Mismatches 65; Indels 12; Gaps 4;

Qy 1 EPIAIVMACRFPFGVTSADDFMDLISSEODALIGFPTDGMWDLDTLYDPDPHPTCYT 60
Db 3332 EPIAIVMACRLPGVTSPEELWELSSGGISGFPTDGMWDLHLHPDPHPTCYT 3391
Qy 61 RNGGFLYDAGHPDAEPFGISPREALAMPPOORLLLETAMETIHAAGINPTLHGTPTGVF 120
Db 3392 RHGGFLHEAGLPAGGFGISPREAVAMPPOORLLLETAMETIHAAGINPTLHGTPTGVF 3451
Qy 121 TGNNGDYALRVHNAAGST-----DGPALTGTAGSVISGRISYTFGREGPAVSVDTACSS 175
Db 3452 TGMFGQGYVA---PDGVSVPTELBEGLAGTSSSVASGRVSYTFGREGPAVSVDTACSS 3507
Qy 176 SLVALHLAQLRAGCSMALAGVTVMSSPGAFAVEFSRQGLAADGHKAPSAADGTG 235
Db 3508 SLVALHLAQLRAGCSMALAGVTVMSSPGAFAVEFSRQGLAADGHKAPSAADGTG 3567
Qy 236 WEGSGMLIVERLSDAHRNGHRLAVVRSAAVQDASNGLTAPNGSQORVIRQALANAGLSAG 295
Db 3568 WEGSGMLIVERLSDAHRNGHRLAVVRSAAVQDASNGLTAPNGSQORVIRQALANAGLSAG 3627
Qy 296 GLSAGVDVAEHAHTGTTGTDPIEAQALLATYQDRAEGPLMLGSVKSNGVHTQAAAGV 355
Db 3628 GLASDVDDVAEHAHTGTTGTDPIEAQALLATYQDRAEGPLMLGSVKSNGVHTQAAAGV 3685
Qy 356 AGVAKWMLRHGLPRTLHVDEPSPHVMSAGAVOLLTETVPMGEGRLRAGVSSFGV 415
Db 3686 AGVAKWMLRHGLPRTLHVDEPSPHVMSAGAVOLLTETVPMGEGRLRAGVSSFGV 3744
Qy 416 VSGTNAHVILEEAP 430
Db 3745 VSGTNAHVILEEAP 3759

RESULT 4

T17410
polyketide synthase type I - Streptomyces venezuelae
C/Species: Streptomyces venezuelae

C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
 C/Accession: T17410
 R/Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
 Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
 A/Title: A gene cluster for macrolide antibiotic biosynthesis in streptomyces venezuelae
 A/Reference number: Z18773; MUID:96445333; PMID:9770448
 A/Accession: T17410
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-3739 <XUE>
 A/Cross-references: EMBL:AF079138; NID:g3808326; PID:g3808035; PIDN:AA69330.1
 C/Genetic:
 A/Gene: pIKAI1
 C/Superfamily: acyl carrier protein homology
 C/Keywords: antibiotic biosynthesis; carrier protein
 F/1445-1516/Domain: acyl carrier protein homology <ACPI>
 F/3570-3641/Domain: acyl carrier protein homology <ACPI>

Query Match 71.9%; Score 1628.5; DB 2; Length 3739;
 Best Local Similarity 72.6%; Pred. No. 3.6e-104;
 Matches 312; Conservative 42; Mismatches 75; Indels 1; Gaps 1;

QY 1 EPIAIVGMACRPFPGVTSADDFWDLISSEBDAIGFPTDRGMDLTLYDPPDPHPTCYT 60
 DB EPIAIVGMACRPLPGVAVSPEDLWRLVAGGDAISGFQDRGMVDEGLYHDPDPHPTSYV 1599
 QY 61 RINGFLVDAGHFDPAEFGISPREALAMDPOORLLLETAMETIBHAGINPHTLHGTPTGVF 120
 DB RINGFLVDAGHFDPAEFGISPREALAMDPOORLLLETAMETIBHAGINPHTLHGTPTGVF 1659
 QY 121 TGTNGODYALRVHNAAGSTDFGALTGTAGSVISGRISYTFEGEPGPAVSDPTACSSSLVAL 180
 DB AGTNGPHYBPLRLNTAEDLEGVGTGNASISGRVSTYGLLEPALTVDPTACSSSLVAL 1719
 QY 1660 AGTNGPHYBPLRLNTAEDLEGVGTGNASISGRVSTYGLLEPALTVDPTACSSSLVAL 1719
 DB 1660 AGTNGPHYBPLRLNTAEDLEGVGTGNASISGRVSTYGLLEPALTVDPTACSSSLVAL 1719
 QY 181 HLAQALRAGECSMALAGVTWSSPGAFVFEFRQGLADGCKAFSAADGTGMEGV 240
 DB HLAQALRAGECSMALAGVTWSSPGAFVFEFRQGLADGCKAFSAADGTGMEGV 1779
 QY 241 GMLLVERLSDAHRNGHVLAVVRSANODGASNGLTAPNGPSOQRYIRALANAGSAG 300
 DB GMLLVERLSDAHRNGHVLAVVRSANODGASNGLTAPNGPSOQRYIRALANAGSAG 1839
 QY 301 DVDAVEAHGTGTLGDPTEAQAALATYGGDRAGEPLMLGSVSNVGHQTQAAGVAVIK 360
 DB DVDAVEAHGTGTLGDPTEAQAALATYGGDRAGEPLMLGSVSNVGHQTQAAGVAVIK 1899
 QY 361 MMALRGLPRLTAVDEBSPHVDMSAGAVQLTETVPW- GEGRLRRAGVSSFGVSGT 419
 DB MMALRGLPRLTAVDEBSPHVDMSAGAVQLTETVPW- GEGRLRRAGVSSFGVSGT 1959
 QY 420 NAHVILEAP 429
 DB NAHVILEAP 1969

RESULT 5
 T17409
 polyketide synthase type I - Streptomyces venezuelae
 C/Species: Streptomyces venezuelae
 C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
 C/Accession: T17409
 R/Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
 Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
 A/Title: A gene cluster for macrolide antibiotic biosynthesis in streptomyces venezuelae
 A/Reference number: Z18773; MUID:96445333; PMID:9770448
 A/Accession: T17409
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-4613 <XUE>
 A/Cross-references: EMBL:AF079138; NID:g3808326; PID:g3808034; PIDN:AA69329.1
 C/Genetic:
 A/Gene: pIKAI
 C/Superfamily: acyl carrier protein homology

C/Keywords: antibiotic biosynthesis; carrier protein
 F/1010-1081/Domain: acyl carrier protein homology <ACPI>
 F/2495-2566/Domain: acyl carrier protein homology <ACPI>
 F/4407-4478/Domain: acyl carrier protein homology <ACPI>

Query Match 71.6%; Score 1622.5; DB 2; Length 4613;
 Best Local Similarity 72.1%; Pred. No. 1.2e-103;
 Matches 310; Conservative 42; Mismatches 77; Indels 1; Gaps 1;

QY 1 EPIAIVGMACRPFPGVTSADDFWDLISSEBDAIGFPTDRGMDLTLYDPPDPHPTCYT 60
 DB EPIAIVGMACRPLPGVAVSPEDLWRLVAGGDAISGFQDRGMVDEGLYHDPDPHPTSYV 1168
 QY 61 RINGFLVDAGHFDPAEFGISPREALAMDPOORLLLETAMETIBHAGINPHTLHGTPTGVF 120
 DB RINGFLVDAGHFDPAEFGISPREALAMDPOORLLLETAMETIBHAGINPHTLHGTPTGVF 1228
 QY 121 TGTNGODYALRVHNAAGSTDFGALTGTAGSVISGRISYTFEGEPGPAVSDPTACSSSLVAL 180
 DB TGTNGODYALRVHNAAGSTDFGALTGTAGSVISGRISYTFEGEPGPAVSDPTACSSSLVAL 1288
 QY 181 HLAQALRAGECSMALAGVTWSSPGAFVFEFRQGLADGCKAFSAADGTGMEGV 240
 DB HLAQALRAGECSMALAGVTWSSPGAFVFEFRQGLADGCKAFSAADGTGMEGV 1348
 QY 241 GMLLVERLSDAHRNGHVLAVVRSANODGASNGLTAPNGPSOQRYIRALANAGSAG 300
 DB GMLLVERLSDAHRNGHVLAVVRSANODGASNGLTAPNGPSOQRYIRALANAGSAG 1408
 QY 301 DVDAVEAHGTGTLGDPTEAQAALATYGGDRAGEPLMLGSVSNVGHQTQAAGVAVIK 360
 DB DVDAVEAHGTGTLGDPTEAQAALATYGGDRAGEPLMLGSVSNVGHQTQAAGVAVIK 1468
 QY 361 MMALRGLPRLTAVDEBSPHVDMSAGAVQLTETVPW- GEGRLRRAGVSSFGVSGT 419
 DB MMALRGLPRLTAVDEBSPHVDMSAGAVQLTETVPW- GEGRLRRAGVSSFGVSGT 1528
 QY 420 NAHVILEAP 429
 DB NAHVILEAP 1538

RESULT 6
 T17463
 rifamycin polyketide synthase module 1-3 - Amycolatopsis mediterranei
 C/Species: Amycolatopsis mediterranei
 C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
 C/Accession: T17463
 R/Schupp, T.
 submitted to the EMBL Data Library, December 1997
 A/Reference number: Z18802
 A/Accession: T17463
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-4735 <SCH>
 A/Cross-references: EMBL:AJ22012; NID:e1227119; PID:e1227120; PIDN:CA11035.1
 A/Experimental source: Strain LBG A3136
 C/Superfamily: acyl carrier protein homology; acetate-CoA ligase homology
 F/53-500/Domain: acyl carrier protein homology <ACPI>
 F/543-610/Domain: acyl carrier protein homology <ACPI>
 F/2102-2173/Domain: acyl carrier protein homology <ACPI>
 F/3079-3150/Domain: acyl carrier protein homology <ACPI>
 F/4578-4649/Domain: acyl carrier protein homology <ACPI>

Query Match 71.1%; Score 1611; DB 2; Length 4735;
 Best Local Similarity 73.3%; Pred. No. 7.9e-103;
 Matches 315; Conservative 33; Mismatches 78; Indels 4; Gaps 3;

QY 1 EPIAIVGMACRPFPGVTSADDFWDLISSEBDAIGFPTDRGMDLTLYDPPDPHPTCYT 60
 DB EPIAIVGMACRPLPGVAVSPEDLWRLVAGGDAISGFQDRGMVDEGLYHDPDPHPTSYV 2254

QY 61 RNCGLYDAGHFDFAFFGSIISPREALAMPDPOQRLLLETAMETIEHAGINPHTLHGTPTGVF 120
D 2255 RHGAFDLDAGFDFAFFGSIISPREALAMPDPOQRLLLETAMETIEHAGINPHTLHGTPTGVF 2314
QY 121 TGNIGDYLALRVNAGOSTDGFALTGTAGSVISGRISTTTFGEEBPASVDTACSSSLVAL 180
D 2315 AGVNSHYSMRMRRAA-GVEGFRLTGGSASVLSGRVAVHFGVEBPATVDTACSSSLVAL 2373
QY 181 HLAQALRAGCSMALAGVTVNMSPGAFFVFSRORGLAADGHCKAFSAADGTGWMG 240
D 2374 HMAVQALRORGCSPALAGVTVNMSPGAFFVFSRORGLAADGHCKAFSAADGTGWMG 2433
QY 241 GMLIVERLSDAHRNGHRLAVVRSASAVNODGASNGLTAPNPGSOORYIRQALANAGISAG 300
D 2434 GLILVERLSEARORGHQVLAIVRGSANVSDASNGLTAPNPGSOORYIRQALANAGISAG 2493
QY 301 DVDAVEAHGTGTTGDPTEAQAALLATTGQDRAGEPPLMGSVKSNGHTQAAAGVAVIK 360
D 2494 DVDAVEAHGTGTTGDPTEAQAALLATTGQDRAGEPPLMGSVKSNGHTQAAAGVAVIK 2551
QY 361 MVMALRHGLPRTLHVDSPHVMMSAGAVQLTETVPWPGEGRLRAGVSSFGVSGTN 420
D 2552 MVMALRHGLPRTLHVDSPHVMMSAGAVQLTETVPWPGEGRLRAGVSSFGVSGTN 2610
QY 421 AHVILEBAPA 430
D 2611 AHVILEBAPA 2620

RESULT 7

117467
rifamycin polyketide synthase modules 9-10 - Amycolatopsis mediterranei
C:Species: Amycolatopsis mediterranei
C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
C:Accession: T17467
R:Schupp, T.
submitted to the EMBL Data Library, December 1997
A:Reference number: Z18802
A:Accession: T17467
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3413 <SCH>
A:Cross-references: EMBL:AJ223012; NID:e1227119; PID:e1227124; PIDN:CAA11039.1
A:Experimental source: strain LBG A3136
C:Superfamily: acyl carrier protein homology
C:Keywords: carrier protein
F:1608-1679/Domain: acyl carrier protein homology <ACP1>
F:3334-3405/Domain: acyl carrier protein homology <ACP2>

Query Match 70.3%; Score 1591.5; DB 2; Length 3413;

Best Local Similarity 72.0%; Pred. No. 1-2e-101; Mismatches 311; Conservative 37; Indels 11; Gaps 4;

QY 1 EPIAIVGACRPFPGVTSADDFMDLISSEDAIGFPTDRGMDLDTLDPDPHPTGCT 60
D 31 EPIAIVGACRPFPGVTSADDFMDLISSEDAIGFPTDRGMDLDTLDPDPHPTGCT 90
QY 61 RNCGLYDAGHFDFAFFGSIISPREALAMPDPOQRLLLETAMETIEHAGINPHTLHGTPTGVF 120
D 91 RNCGLYDAGHFDFAFFGSIISPREALAMPDPOQRLLLETAMETIEHAGINPHTLHGTPTGVF 150
QY 121 TGNIGDYLALRVNAGOSTDGFALTGTAGSVISGRISTTTFGEEBPASVDTACSSSLVAL 177
D 151 TGNIGDYLALRVNAGOSTDGFALTGTAGSVISGRISTTTFGEEBPASVDTACSSSLVAL 205
QY 178 HLAQALRAGCSMALAGVTVNMSPGAFFVFSRORGLAADGHCKAFSAADGTGWMG 237
D 206 HLAQALRAGCSMALAGVTVNMSPGAFFVFSRORGLAADGHCKAFSAADGTGWMG 265
QY 238 EGVMGLVERLSDAHRNGHRLAVVRSASAVNODGASNGLTAPNPGSOORYIRQALANAGISAG 297
D 266 EGVMGLVERLSDAHRNGHRLAVVRSASAVNODGASNGLTAPNPGSOORYIRQALANAGISAG 325

QY 298 SAGDVDAVEAHGTGTTGDPTEAQAALLATTGQDRAGEPPLMGSVKSNGHTQAAAGVAVIK 357
D 326 SAGDVDAVEAHGTGTTGDPTEAQAALLATTGQDRAGEPPLMGSVKSNGHTQAAAGVAVIK 383
QY 358 VIRMVQALRHGAMPRTLHVAEPTEVDWMSAGAVELLTEPRAWPADG-RPRRAGVSAFGIS 417
D 384 VIRMVQALRHGAMPRTLHVAEPTEVDWMSAGAVELLTEPRAWPADG-RPRRAGVSAFGIS 442
QY 418 GTNAHVLIEBAP 429
D 443 GTNAHVLIEBAP 454

RESULT 8

117428
FK506 polyketide synthase - Streptomyces sp. (strain MA6548)
C:Species: Streptomyces sp.
A:Variety: strain MA6548
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000
C:Accession: T17428
R:McLarni, H.; Shafiee, A.
Eur. J. Biochem. 256, 528-534, 1998
A:Title: The biosynthetic gene cluster for the macrolactone ring of the immunosuppressor
A:Reference number: Z18779; M01D:98451508; PMID:9780228
A:Accession: T17428
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7576 <MOT>
A:Cross-references: EMBL:AF082100; NID:93798623; PID:93798624; PIDN:AAC68815.1
A:Experimental source: strain MA6548
C:Genetics:
A:Gene: fkbB
C:Function:
A:Description: involved in synthesis of the backbone of the immunosuppressor FK506 polyketide
C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase
C:Keywords: carrier protein
F:54-500/Domain: acetate-CoA ligase homology <ACL>
F:1095-1166/Domain: acyl carrier protein homology <ACP1>
F:1204-1599/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F:1682-1953/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F:2680-2751/Domain: acyl carrier protein homology <ACP2>
F:2804-3198/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F:3295-3569/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F:4320-4391/Domain: acyl carrier protein homology <ACP3>
F:4435-4830/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F:5003-5974/Domain: acyl carrier protein homology <ACP4>
F:6018-6412/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F:6513-6785/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
F:7450-7521/Domain: acyl carrier protein homology <ACP5>

Query Match 70.2%; Score 1590; DB 2; Length 7576;

Best Local Similarity 73.9%; Pred. No. 4e-101; Mismatches 315; Conservative 37; Indels 8; Gaps 3;

QY 1 EPIAIVGACRPFPGVTSADDFMDLISSEDAIGFPTDRGMDLDTLDPDPHPTGCT 60
D 1183 EPIAIVGACRPFPGVTSADDFMDLISSEDAIGFPTDRGMDLDTLDPDPHPTGCT 1242
QY 61 RNCGLYDAGHFDFAFFGSIISPREALAMPDPOQRLLLETAMETIEHAGINPHTLHGTPTGVF 120
D 1243 RNCGLYDAGHFDFAFFGSIISPREALAMPDPOQRLLLETAMETIEHAGINPHTLHGTPTGVF 1302
QY 121 TGNIGDYLALRVNAGOSTDGFALTGTAGSVISGRISTTTFGEEBPASVDTACSSSLVAL 180
D 1303 TGNIGDYLALRVNAGOSTDGFALTGTAGSVISGRISTTTFGEEBPASVDTACSSSLVAL 1357
QY 181 HLAQALRAGCSMALAGVTVNMSPGAFFVFSRORGLAADGHCKAFSAADGTGWMG 240
D 1358 HLAQALRAGCSMALAGVTVNMSPGAFFVFSRORGLAADGHCKAFSAADGTGWMG 1417
QY 241 GMLIVERLSDAHRNGHRLAVVRSASAVNODGASNGLTAPNPGSOORYIRQALANAGISAG 300
D 1418 GMLIVERLSDAHRNGHRLAVVRSASAVNODGASNGLTAPNPGSOORYIRQALANAGISAG 1477

Qy	301	VVDVAEAAAGTGTTLGDDPIEAOALATYGGDRAGEPLMGSKSVNGHTQAAAGVAGYIK	360
Db	1478	DVDAVEAAGTGTVLDDPIEAQVNLVSYYGER--EYPLLGSKSVNGHTQAAAGVAGYIK	1533
Qy	361	NMMALRHGILLPTTLHVDESPPHVMDSAGAVOLLTETVPMGEGEGRLLRAGVSSFCVSGTN	420
Db	1536	NMMARHGILLPTTLHVDESPSHVMDSAGAVEVLTETRPM-ESGARPRAGVSSFCVSGTN	1594
Qy	421	AHVILE 426	
Db	1595	AHVILE 1600	

RESULT 9

polyketide synthase - Streptomyces hygroscopicus
C.Species: Streptomyces hygroscopicus
C.Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
C.Accession: J30225
R.Aparticito, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Staun,
Gene 169, 9-16, 1996
A.Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hygroscopicus
A.Reference number: Z20782; MUID:96186896; PMID:8635756
A.Accession: J30225
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-10223 <APA>
A.Cross-references: EMBL:X86780; NID:g987008; PID:g987099; PIDN:CA60459.1
A.Genetic: map
C.Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protein
C.Keywords: carrier protein
F.554.449/Domain: 3-oxoacyl-[acyl-carrier-protein] homology <OAS1>
F.1533-1604/Domain: acyl carrier protein homology <ACP>
F.1533-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F.1267-2409/Domain: acyl carrier protein homology <ACP>
F.3132-3243/Domain: acyl carrier protein homology <ACP>
F.3287-3681/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F.3778-4052/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F.5160-5231/Domain: acyl carrier protein homology <ACP>
F.5275-5667/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F.6776-6847/Domain: acyl carrier protein homology <ACP>
F.6891-7285/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS5>
F.8411-8482/Domain: acyl carrier protein homology <ACP>
F.8526-8681/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS6>
F.9102-9285/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
F.10069-10140/Domain: acyl carrier protein homology <ACP>

Query Match	Score	DB 2;	Length
70.0%	1586;	DB 2;	10223;

Matches 312; Conservative 34; Mismatches 75; Indels 8; Gaps 3;

Matches 312; Conservative 34; Mismatches 75; Indels 8; Gaps 3;

QY 1 BEIALVGNACRPGCVTSADDDFWDIISSSQDAIGCFPTRGMDLTLYLDPDPDHGTCYT 60
 Db 33 BFLAIVGNACRIRPGCVSSPEDLMRLVESTDAISGFPADRGMDVENLYLDPDEAAGKSYC 92
 QY 61 RNGEFLYDAGHFDAPFPGISPREALAMPDQORLLLETAMETTERAGINPHTLHGTPGVF 120
 Db 93 VGGFLIDSGGFDAPFPGISPREALAMPDQORLYVEASWEAEPRRGIIEGYSIRGSDTVGF 152
 QY 121 TGTNGODYALRNVHNAQGSTDGFALVTGTGVSISGRISYTFGEGGAVSVDTACSSLYAL 180
 Db 153 MCAVYFGCV-----GADLGGFGALTAGAVSVLSGKVSYTFGLGGAVYVDTACSSLYAL 207
 QY 161 HLAICALNPAEGCSMALAGVTVMSSPGAFVEFVSROKGLAADGHCKAFAAADGTWGEQV 240
 Db 208 HDAAYALRQEGECSLILVGGVTVMATPQGSFVEFSRORGLASDRCACAFADSADGTWMAEGA 267
 QY 241 GMLTVERISDARNRHRLVAVVRGSANNDGASNGLTLPNGSQQORVIRQALANNGLSAG 300
 Db 268 GVLTVERSLSDAQRKGIQVLAIVVRGSANNDGASNGLSAENGSSQORVIRALASNNGLSTA 327

```

0Y      DVAVEAHGCTGTLGPPIAQAQLLATTYGDDRAGEBPLMGSVKSVNGHTOAAAGVAGIK 360
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      EVDVEAHGCTGTLGPIAQAQLLATTYGDDR--BEPPLMGSVKSNLGHTOAAAGVAGIYK 380
0Y      MMALRHGLLPTLHVDDESPHVDMSACAVOLLTETVPMGCGEGRLRAGVSSFVGSGTN 420
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      MMALHGLHVLPTLHVDDESPRHVDMDTDAVALVTNQMPP--DMGRPRAGVSSFVGISGTN 440
0Y      AHVILEEAP 429
      |||||F||
Db      AHVILEEAP 453

```

RESULT 10

polyketide synthase - Streptomyces hygroscopicus
C:Species: Streptomyces hygroscopicus
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 01-Dec-2000
R:Accession: J30226
R:Apertio, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Stan-
Gene 169, 9-16, 1996
A:Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hy-
A:Reference number: Z20782; MUID:96186896; PMID:865756
A:Accession: J30226
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-8563 <APA>
A:Cross-references: EMBL:X68780; NID:J987088; PID:J987100; PIDN:CAA60460.1
C:Genetics:
A:Gene: rapA
C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:54-503/Domain: acetate-CoA ligase homology <ACL>
F:1329-1724/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F:1817-2091/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F:3199-3270/Domain: acyl carrier protein homology <ACP1>
F:3314-3706/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F:4787-4858/Domain: acyl carrier protein homology <ACP2>
F:4992-5293/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F:5386-5659/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F:5760-6831/Domain: acyl carrier protein homology <ACP3>
F:6875-7269/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F:7362-7636/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
F:8412-8483/Domain: acyl carrier protein homology <ACP4>
F:14822/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match	69.8%;	Score 1581.5;	DB 2;	Length 8563
-------------	--------	---------------	-------	-------------

Best Local Similarity 1/2.04; P-Val. NO. 1.88-100;
Matches 309; Conservative 40; Mismatches 71; Indels 9; Gaps 4;

Best Local Similarity 1/2.04; P-Val. NO. 1.88-100;
Matches 309; Conservative 40; Mismatches 71; Indels 9; Gaps 4;

QY	1	EPRIAVGNACFPFGVTSADFMFLISSBOALICGPTDGMJDLDTYVDPDHPGTCYTT	60
Db	6654	EPRLVGNACRLPGGVSSPEDLMRLVSGTDAVSGFPTDGMQDVENLYDSDPEAAKSYC	69133
QY	61	RNGSFLLDAGHFDAEFPCISPRELLANDPOORLLLETAWEIIEHAGINPHRLHGTPTGVF	120
Db	6914	VQSGFLLPTTAGFDGPFGISPRELLANDPOORLLLEYSWEAFEEAGIEPGSVRSDDGVF	69737
QY	121	TGTNGODYALRVHNAAGOSTDGFALTGTAGSAVYSGRISYTFEGFEGPAASVDPACSSSLVAL	180
Db	6974	IGAPFVGY-----GAGFDREGYCAT--SGPSVLSGRSAVYVGLBEPALITMDTACSSSLVAL	7027
QY	181	HLACQALRAGECCSMALLAGVYTWSSPCAFAPEFSQORGLANDGHKKAASAAADGTGMKEGV	240
Db	7028	HLAAQALRNGCSSLALLAGVYVMAATPEVFEFAQRLGASGRKAAKADASDAGGFSEGA	70877
QY	241	GMLIVERTSDAHRNGHRLVAVNRGSANODGASNGLTAPNGPSQOARYIROLAANAGLSAG	300
Db	7088	GLLIVERTSDARRRNGHQLVAVNRGSANVODGASNGLTAPNGPSSQORYIIPALASNAIGLSTA	7147
QY	301	DVDAVEAHGCTTYIGDPIEAQALLATTGODRAGEGPIMLGSVYSNVGHDTQAAGVAGYIK	360
Db	7148	DVDVVEAHGCTTYIGDPIEAQALLATTGQOR--EOPILLGSLKNIGHTQAAGSVAGYIK	7205

Query Match 69.2%; Score 1567.5; DB 2; Length 1763;
Best Local Similarity 71.1%; Pred. No. 2.4e-100;
Matches 305; Conservative 37; Mismatches 82; Indels 5; Gaps 3;

QY 1 EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 60
DB EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 60
QY 36 EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 95
DB EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 95
QY 61 RNSGFLVDAGHFDPAEFPGISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 120
DB RNSGFLVDAGHFDPAEFPGISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 120
QY 96 DGGFPLHDALFDGFGFISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 155
DB DGGFPLHDALFDGFGFISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 155
QY 121 TGTNGODYALRVHNAAGOSTDGFALTGTAGVISGRISYTFEGEPGPAVSVDPAACSSSLVAL 180
DB TGTNGODYALRVHNAAGOSTDGFALTGTAGVISGRISYTFEGEPGPAVSVDPAACSSSLVAL 180
QY 156 TGAAGSGVGGGL--TGPEMOSFAGTGLASSVASGRVSVFPEEPATYIDTACSSSLVAM 213
DB TGAAGSGVGGGL--TGPEMOSFAGTGLASSVASGRVSVFPEEPATYIDTACSSSLVAM 213
QY 181 HLACQALRAGECSMALAGGVTVMSFGAFVPSRQGLADGHCAPSAADGTGMBGV 240
DB HLACQALRAGECSMALAGGVTVMSFGAFVPSRQGLADGHCAPSAADGTGMBGV 240
QY 214 HLAAQALRQDCSNALAGGAMVMSGPDSPFVFSRQGLATDGRKAPASGADGAVLAEGI 273
DB HLAAQALRQDCSNALAGGAMVMSGPDSPFVFSRQGLATDGRKAPASGADGAVLAEGI 273
QY 241 GMLVERLSDAHNRGHRVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 300
DB GMLVERLSDAHNRGHRVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 300
QY 274 SVVVERLSVARERGHVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 333
DB SVVVERLSVARERGHVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 333
QY 301 DVDVAEAGTGTTLGDPDIEAQAALLATYGGDRAGEGPMILGSKNVNCHTQAAAGVAVIK 360
DB DVDVAEAGTGTTLGDPDIEAQAALLATYGGDRAGEGPMILGSKNVNCHTQAAAGVAVIK 360
QY 334 DVDVEAHGTGTSLGDPDIEAQAALLATYGGDR--ETPLMGLSKNIGHTQAAAGVAVIK 391
DB DVDVEAHGTGTSLGDPDIEAQAALLATYGGDR--ETPLMGLSKNIGHTQAAAGVAVIK 391
QY 361 MMALRGLPRLTHVDEPSPHVDMSAGAVOLLTEVPWPGEGRLRAGVSVFVSGTN 420
DB MMALRGLPRLTHVDEPSPHVDMSAGAVOLLTEVPWPGEGRLRAGVSVFVSGTN 420
QY 392 VQALRRGIVMPRLTHVDEPSPHVDMSAGAVOLLTEVPWPGEGRLRAGVSVFVSGTN 450
DB VQALRRGIVMPRLTHVDEPSPHVDMSAGAVOLLTEVPWPGEGRLRAGVSVFVSGTN 450
QY 421 AHVLEEAR 429
DB AHVLEEAR 429
QY 451 VHLIIEBP 459
DB VHLIIEBP 459

RESULT 14

T03224
probable polyketide synthase module 4 - Streptomyces hygroscopicus
C/Specties: Streptomyces hygroscopicus
C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 26-May-2000
C/Accession: T03224
R/Ruan, X.; Stasbel, D.; Lax, S.; Katz, L.
Gene 203, 1-9, 1997
A/Title: A second type-1 PKS gene cluster isolated from Streptomyces hygroscopicus ATCC
A/Reference number: Z14848; MUID:98085969; PMID:9426000
A/Accession: T03224
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1937 <RUA>
A/Cross-references: EMBL:AF007101; NID:g2624946; PIDN:ACG38064.1; PID:g2624951
A/Experimental source: ATCC 29253
C/Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoac
homology; [acyl-carrier-protein] S-malonyltransferase homology
C/Keywords: carrier protein
F/549-827/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>

Query Match 69.0%; Score 1563; DB 2; Length 1937;
Best Local Similarity 69.5%; Pred. No. 5.5e-100;
Matches 296; Conservative 49; Mismatches 80; Indels 2; Gaps 1;

QY 1 EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 60
DB EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 60
QY 34 EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 93
DB EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 93
QY 61 RNSGFLVDAGHFDPAEFPGISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 120
DB RNSGFLVDAGHFDPAEFPGISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 120
QY 94 RDSGFLVDAGHFDPAEFPGISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 153
DB RDSGFLVDAGHFDPAEFPGISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 153
QY 121 TGTNGODYALRVHNAAGOSTDGFALTGTAGVISGRISYTFEGEPGPAVSVDPAACSSSLVAL 180
DB TGTNGODYALRVHNAAGOSTDGFALTGTAGVISGRISYTFEGEPGPAVSVDPAACSSSLVAL 180

DB 154 TGVHNDYADMSADDTDLQLEGVLNGSAAVASGRVSYTFEGEPGPAVSVDPAACSSSLVAL 213
DB 181 HLACQALRAGECSMALAGGVTVMSFGAFVPSRQGLADGHCAPSAADGTGMBGV 240
DB 214 HLAAQALRQDCSNALAGGAMVMSGPDSPFVFSRQGLATDGRKAPASGADGAVLAEGI 273
QY 241 GMLVERLSDAHNRGHRVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 300
DB GMLVERLSDAHNRGHRVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 300
QY 274 SVVVERLSVARERGHVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 333
DB SVVVERLSVARERGHVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 333
QY 301 DVDVAEAGTGTTLGDPDIEAQAALLATYGGDRAGEGPMILGSKNVNCHTQAAAGVAVIK 360
DB DVDVAEAGTGTTLGDPDIEAQAALLATYGGDRAGEGPMILGSKNVNCHTQAAAGVAVIK 360
QY 334 DVDVEAHGTGTSLGDPDIEAQAALLATYGGDR--ETPLMGLSKNIGHTQAAAGVAVIK 391
DB DVDVEAHGTGTSLGDPDIEAQAALLATYGGDR--ETPLMGLSKNIGHTQAAAGVAVIK 391
QY 361 MMALRGLPRLTHVDEPSPHVDMSAGAVOLLTEVPWPGEGRLRAGVSVFVSGTN 420
DB MMALRGLPRLTHVDEPSPHVDMSAGAVOLLTEVPWPGEGRLRAGVSVFVSGTN 420
QY 392 VQALRRGIVMPRLTHVDEPSPHVDMSAGAVOLLTEVPWPGEGRLRAGVSVFVSGTN 451
DB VQALRRGIVMPRLTHVDEPSPHVDMSAGAVOLLTEVPWPGEGRLRAGVSVFVSGTN 451
QY 421 AHVLEEAR 429
DB AHVLEEAR 429
QY 452 VHLIIEBP 460
DB VHLIIEBP 460

RESULT 15

T17411
polyketide synthase III - Streptomyces venezuelae
C/Specties: Streptomyces venezuelae
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 03-Nov-2000
C/Accession: T17411
R/Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
A/Title: A gene cluster for macroide antibiotic biosynthesis in streptomyces venezuela
A/Reference number: Z18773; MUID:98445333; PMID:9770448
A/Accession: T17411
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1562 <XUE>
A/Cross-references: EMBL:AF079138; NID:g3808326; PID:g3800836; PIDN:ACG9331.1
C/Genetics:
A/Gene: p1X411
C/Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoac
homology; [acyl-carrier-protein] S-malonyltransferase homology
C/Keywords: antibiotic biosynthesis; carrier protein
F/1403-1474/Domain: acyl carrier protein homology <AC>

Query Match 68.3%; Score 1561.5; DB 2; Length 1562;
Best Local Similarity 69.8%; Pred. No. 5.4e-100;
Matches 305; Conservative 47; Mismatches 74; Indels 11; Gaps 3;

QY 1 EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 60
DB EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 60
QY 35 EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 94
DB EPIAIVGMACRFPGCVTSADFPMDLISSEDAIGFPPTDRGMDLDTLYDPDPHPTCYT 94
QY 61 RNSGFLVDAGHFDPAEFPGISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 120
DB RNSGFLVDAGHFDPAEFPGISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 120
QY 95 RSGFPLHDAGHFDPAEFPGISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 154
DB RSGFPLHDAGHFDPAEFPGISPREBALMDPOORLLLETAMETIHAGINPHLTGTPGVF 154
QY 121 TG-----TNGODYALRVHNAAGOSTDGFALTGTAGVISGRISYTFEGEPGPAVSVDPAAC 174
DB TG-----TNGODYALRVHNAAGOSTDGFALTGTAGVISGRISYTFEGEPGPAVSVDPAAC 174
QY 155 VGMHNGYTGQTAAVQ---SPELEGHVLSGALGLSLGRIVVLTGTDGPAALTVDPAC 210
DB VGMHNGYTGQTAAVQ---SPELEGHVLSGALGLSLGRIVVLTGTDGPAALTVDPAC 210
QY 175 SSIVLAHLAQAALRAGECSMALAGGVTVMSFGAFVPSRQGLADGHCAPSAADGT 234
DB SSIVLAHLAQAALRAGECSMALAGGVTVMSFGAFVPSRQGLADGHCAPSAADGT 234
QY 235 GMEGEGVGLVERLSDAHNRGHRVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 294
DB GMEGEGVGLVERLSDAHNRGHRVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 294
QY 271 GPMAGGVLLVERLSDAHNRGHRVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 330
DB GPMAGGVLLVERLSDAHNRGHRVLAIVVRSANOGASNGLTAPNGPSQORVYRQALNAGLSAG 330
QY 295 AGLSAGVDVAEAGTGTTLGDPDIEAQAALLATYGGDRAGEGPMILGSKNVNCHTQAAAG 354
DB AGLSAGVDVAEAGTGTTLGDPDIEAQAALLATYGGDRAGEGPMILGSKNVNCHTQAAAG 354
QY 331 ARLAPGVDDVAEAGTGTTLGDPDIEAQAALLATYGGDRAGEGPMILGSKNVNCHTQAAAG 390
DB ARLAPGVDDVAEAGTGTTLGDPDIEAQAALLATYGGDRAGEGPMILGSKNVNCHTQAAAG 390

QY 355 VAGVIRKMMALRHGILLPRTLHVDEBSPHYDWSAGAVOLLTETVPMP-CGEGRLRRAGVSS 413
 Db 391 VAGVIRKMMALRHGILLPRTLHVDEBSPHYDWSAGAVOLLTETVPMP-CGEGRLRRAGVSS 450
 QY 414 FGVSQTNAHVILLEAPA 430
 Db 451 FGVSQTNAHVILLEAPA 467

Search completed: June 17, 2003, 13:12:23
 Job time : 10.6019 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 13:08:02 ; Search time 13.398 Seconds
(without alignments)
3431.399 Million cell updates/sec

Title: US-09-914-286-3_COPY_2086_2515

Perfect score: 2265 1 EPIAIVGMACRPFPGVTSAD.....VSSFGVSGTNAYHLEAPA 430

Sequences:

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/1/pubppaa/US08 NEW PUB. pep.*
2: /cgn2_6/ptodata/1/pubppaa/PCT NEW PUB. pep.*
3: /cgn2_6/ptodata/1/pubppaa/US06 NEW PUB. pep.*
4: /cgn2_6/ptodata/1/pubppaa/US07 NEW PUB. pep.*
5: /cgn2_6/ptodata/1/pubppaa/US07 PUBCOMB. pep.*
6: /cgn2_6/ptodata/1/pubppaa/US07 PUBCOMB. pep.*
7: /cgn2_6/ptodata/1/pubppaa/US08 PUBCOMB. pep.*
8: /cgn2_6/ptodata/1/pubppaa/US08 PUBCOMB. pep.*
9: /cgn2_6/ptodata/1/pubppaa/US09 NEW PUB. pep.*
10: /cgn2_6/ptodata/1/pubppaa/US09 PUBCOMB. pep.*
11: /cgn2_6/ptodata/1/pubppaa/US10 NEW PUB. pep.*
12: /cgn2_6/ptodata/1/pubppaa/US10 PUBCOMB. pep.*
13: /cgn2_6/ptodata/1/pubppaa/US60 NEW PUB. pep.*
14: /cgn2_6/ptodata/1/pubppaa/US60 PUBCOMB. pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1654	72.0	3519	9 US-09-808-880-4	Sequence 4, Appl1
2	1646.5	72.7	4150	9 US-09-808-880-2	Sequence 2, Appl1
3	1628.5	71.9	3739	9 US-09-860-846-33	Sequence 33, Appl1
4	1628.5	71.9	3739	9 US-09-988-384B-33	Sequence 33, Appl1
5	1628.5	71.9	3739	9 US-09-836-821-33	Sequence 33, Appl1
6	1628.5	71.9	3739	9 US-09-793-708-2	Sequence 2, Appl1
7	1628.5	71.9	3739	10 US-09-861-289-33	Sequence 33, Appl1
8	1628.5	71.9	11877	9 US-09-860-846-6	Sequence 6, Appl1
9	1628.5	71.9	11877	9 US-09-836-821-6	Sequence 6, Appl1
10	1628.5	71.9	11877	10 US-09-861-289-6	Sequence 6, Appl1
11	1628.5	71.9	12199	9 US-09-988-384B-6	Sequence 6, Appl1
12	1628.5	71.9	4551	9 US-09-793-708-1	Sequence 1, Appl1
13	1622.5	71.6	4613	9 US-09-860-846-31	Sequence 31, Appl1
14	1622.5	71.6	4613	9 US-09-988-384B-31	Sequence 31, Appl1
15	1622.5	71.6	4613	9 US-09-836-821-31	Sequence 31, Appl1
16	1622.5	71.6	4613	10 US-09-861-289-31	Sequence 31, Appl1
17	1609	69.3	3816	9 US-09-808-880-3	Sequence 3, Appl1
18	1570	69.3	5215	9 US-09-860-846-2	Sequence 2, Appl1
19	1570	69.3	5215	9 US-09-988-384B-2	Sequence 2, Appl1

20	1570	69.3	5215	9 US-09-836-821-2	Sequence 2, Appl1
21	1570	69.3	5215	10 US-09-861-289-2	Sequence 2, Appl1
22	1561.5	68.9	1562	9 US-09-860-846-35	Sequence 35, Appl1
23	1561.5	68.9	1562	9 US-09-988-384B-35	Sequence 35, Appl1
24	1561.5	68.9	1562	9 US-09-836-821-35	Sequence 35, Appl1
25	1561.5	68.9	1562	9 US-09-793-708-3	Sequence 35, Appl1
26	1561.5	68.9	1562	10 US-09-861-289-35	Sequence 35, Appl1
27	1543	68.1	1346	9 US-09-793-708-4	Sequence 37, Appl1
28	1538	67.9	1346	9 US-09-860-846-37	Sequence 37, Appl1
29	1538	67.9	1346	9 US-09-988-384B-37	Sequence 37, Appl1
30	1538	67.9	1346	9 US-09-836-821-37	Sequence 37, Appl1
31	1538	67.9	1346	10 US-09-861-289-37	Sequence 37, Appl1
32	1382.5	61.0	7257	9 US-10-014-717-5	Sequence 5, Appl1
33	1293	57.1	458	10 US-09-892-870-4	Sequence 4, Appl1
34	1180.5	52.1	3798	9 US-10-014-717-6	Sequence 6, Appl1
35	1178	52.0	1421	9 US-10-014-717-2	Sequence 2, Appl1
36	1139.5	50.3	1837	9 US-09-712-363-261	Sequence 261, App
37	1085.5	47.9	2439	9 US-10-014-717-7	Sequence 7, Appl1
38	977.5	43.2	1402	9 US-09-712-363-166	Sequence 166, App
39	942	41.6	1610	9 US-09-738-626-6666	Sequence 6666, Ap
40	887	39.2	1832	10 US-09-775-938A-38	Sequence 38, Appl1
41	860	38.0	1832	9 US-10-014-717-4	Sequence 4, Appl1
42	848	37.4	3032	9 US-09-836-705-44	Sequence 44, Appl1
43	783	34.6	2563	9 US-09-836-705-46	Sequence 46, Appl1
44	715	31.6	2509	9 US-10-237-271-1	Sequence 1, Appl1
45	665	29.4	2756	9 US-10-331-061-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-09-808-880-4
Sequence 4, Application US/09808880
Publication No. US20030027287A1
GENERAL INFORMATION:
APPLICANT: Beclach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/808, 880
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: US/09/428, 517
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/120, 254
PRIOR FILING DATE: 1999-02-16
PRIOR APPLICATION NUMBER: 60/106, 100
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 3519
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: Oleandolide PKS
US-09-808-880-4

Query Match 73.0%; Score 1654; DB 9; Length 3519;
Best Local Similarity 73.3%; Pred. No. 2.2e-121;
Matches 315; Conservative 40; Mismatches 75; Indels 0; Gaps 0;
QY 1 EPIAIVGMACRPFPGVTSADFDWDLISSEDAIGFPTRGMDLTLXPDPDPHPTGCTY 60
Db 1687 EALAIAMSCRFEGGIDSPEDLWFLAEGSDAVAGLPEBRGMOLDLALYHPDPENPTGVV 1746
QY 61 RNCGLYDACHFAEFEGGISPREALAMDPOQRLTLTANETTIBHAGINPTLHGTPTGVF 120
Db 1747 REGAFYDYDAQFDAGGFGISPREALAMDPOQRLTLTANETTIBHAGINPTLHGTPTGVF 1806

QY 121 TGTNGODYALRVHNAAGSTDFALTGTAGSVISGRISYTFGEFEGPAVSVDTACSSSLVAL 180
DB 1807 IGAGHGGGPDPKAPESVAVYLLTGTSAVLSGRISYTFGEFEGPAVSVDTACSSSLVAL 1866
QY 181 HLAQALRAGECSMALAGVTVMSPGAFVFSRQGLADGHCKAFSAADGTGMEGV 240
DB 1867 HLAQALRAGECSMALAGVTVMSPGAFVFSRQGLADGHCKAFSAADGTGMEGV 1926
QY 241 GMLVRLSDAHRNGHRLVAVRGSANVODGASNGLTAPNGPSOQRYRQALANAGISAG 300
DB 1927 SLTLRLSDARRRGRHRLVAVRGSANVODGASNGLTAPNGPSOQRYRQALANAGISAG 1986
QY 301 DVDAVEAHGTGTLGDPPIEAQALATYQDRAGEGPTLGLSVKSNVGHQTAAGAVAGIK 360
DB 1987 DVDAVEAHGTGTLGDPPIEAQALATYQDRAGEGPTLGLSVKSNVGHQTAAGAVAGIK 2046
QY 361 MVALRGLPLRTHLHDEBSPHYDMSAGAVOLLTETVPMFGEGRLRRAVGSFGVSGTN 420
DB 2047 MVALRGLPLRTHLHDEBSPHYDMSAGAVOLLTETVPMFGEGRLRRAVGSFGVSGTN 2106
QY 421 AHVLEERPA 430
DB 2107 AHVLEERPA 2116

RESULT 2

US-09-808-880-2
Sequence 2, Application US/09808880
Publication No. US20030027287A1
GENERAL INFORMATION:
APPLICANT: Beliaich, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/808, 880
PRIORITY FILING DATE: 2001-03-14
PRIORITY APPLICATION NUMBER: US/09/428, 517
PRIORITY FILING DATE: 1999-10-28
PRIORITY APPLICATION NUMBER: 60/120, 254
PRIORITY FILING DATE: 1999-02-16
PRIORITY APPLICATION NUMBER: 60/106, 100
PRIORITY FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4150
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: Oleandolide PKS
US-09-808-880-2

Query Match 72.7%; Score 1646.5; DB 9; Length 4150;
Best Local Similarity 72.8%; Pred. No. 1.1e-120;
Matches 313; Conservative 41; Mismatches 75; Indels 1; Gaps 1;

QY 1 EPIAIVGACRFPGVTSADDFMDLISSEODAIIGFPTDRGMDLDTLYDPDPHPTCYT 60
DB 2546 EPIAIVGACRFPGVTSADDFMDLISSEODAIIGFPTDRGMDLDTLYDPDPHPTCYT 2605
QY 61 RINGFLVDAGHFDPAEFPGISPREALAMPQORLLLETAMETIEHAGINPHLTGTPGVF 120
DB 2606 RINGFLVDAGHFDPAEFPGISPREALAMPQORLLLETAMETIEHAGINPHLTGTPGVF 2665
QY 121 TGTNGODYALRVHNAAGSTDFALTGTAGSVISGRISYTFGEFEGPAVSVDTACSSSLVAL 180
DB 2666 TGTNGODYALRVHNAAGSTDFALTGTAGSVISGRISYTFGEFEGPAVSVDTACSSSLVAL 2725
QY 181 HLAQALRAGECSMALAGVTVMSPGAFVFSRQGLADGHCKAFSAADGTGMEGV 240
DB 1907 HLAQALRAGECSMALAGVTVMSPGAFVFSRQGLADGHCKAFSAADGTGMEGV 2466

DB 2726 HLAQALRAGECSMALAGVTVMSPGAFVFSRQGLADGHCKAFSAADGTGMEGV 2785
QY 241 GMLVRLSDAHRNGHRLVAVRGSANVODGASNGLTAPNGPSOQRYRQALANAGISAG 300
DB 2786 GMLVRLSDAHRNGHRLVAVRGSANVODGASNGLTAPNGPSOQRYRQALANAGISAG 2845
QY 301 DVDAVEAHGTGTLGDPPIEAQALATYQDRAGEGPTLGLSVKSNVGHQTAAGAVAGIK 360
DB 2846 DVDAVEAHGTGTLGDPPIEAQALATYQDRAGEGPTLGLSVKSNVGHQTAAGAVAGIK 2905
QY 361 MVALRGLPLRTHLHDEBSPHYDMSAGAVOLLTETVPMFGEGRLRRAVGSFGVSGTN 420
DB 2906 MVALRGLPLRTHLHDEBSPHYDMSAGAVOLLTETVPMFGEGRLRRAVGSFGVSGTN 2964
QY 421 AHVLEERPA 430
DB 2965 AHVLEERPA 2974

RESULT 3

US-09-860-846-33
Sequence 33, Application US/09860846
Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438051
CURRENT APPLICATION NUMBER: US/09/860, 846
PRIORITY FILING DATE: 2001-05-18
PRIORITY APPLICATION NUMBER: 09/105, 537
PRIORITY FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-860-846-33

Query Match 71.9%; Score 1628.5; DB 9; Length 3739;
Best Local Similarity 72.6%; Pred. No. 2.5e-119;
Matches 312; Conservative 42; Mismatches 75; Indels 1; Gaps 1;

QY 1 EPIAIVGACRFPGVTSADDFMDLISSEODAIIGFPTDRGMDLDTLYDPDPHPTCYT 60
DB 1540 EPIAIVGACRFPGVTSADDFMDLISSEODAIIGFPTDRGMDLDTLYDPDPHPTCYT 1599
QY 61 RINGFLVDAGHFDPAEFPGISPREALAMPQORLLLETAMETIEHAGINPHLTGTPGVF 120
DB 1600 RINGFLVDAGHFDPAEFPGISPREALAMPQORLLLETAMETIEHAGINPHLTGTPGVF 1659
QY 121 TGTNGODYALRVHNAAGSTDFALTGTAGSVISGRISYTFGEFEGPAVSVDTACSSSLVAL 180
DB 1660 TGTNGODYALRVHNAAGSTDFALTGTAGSVISGRISYTFGEFEGPAVSVDTACSSSLVAL 1719
QY 181 HLAQALRAGECSMALAGVTVMSPGAFVFSRQGLADGHCKAFSAADGTGMEGV 240
DB 1720 HLAQALRAGECSMALAGVTVMSPGAFVFSRQGLADGHCKAFSAADGTGMEGV 1779
QY 241 GMLVRLSDAHRNGHRLVAVRGSANVODGASNGLTAPNGPSOQRYRQALANAGISAG 300
DB 1780 GMLVRLSDAHRNGHRLVAVRGSANVODGASNGLTAPNGPSOQRYRQALANAGISAG 1839
QY 301 DVDAVEAHGTGTLGDPPIEAQALATYQDRAGEGPTLGLSVKSNVGHQTAAGAVAGIK 360
DB 1840 DVDAVEAHGTGTLGDPPIEAQALATYQDRAGEGPTLGLSVKSNVGHQTAAGAVAGIK 1899
QY 361 MVALRGLPLRTHLHDEBSPHYDMSAGAVOLLTETVPMFGEGRLRRAVGSFGVSGTN 420
DB 1900 MVALRGLPLRTHLHDEBSPHYDMSAGAVOLLTETVPMFGEGRLRRAVGSFGVSGTN 1959

QY 420 NAHVILEAP 429
 1960 NAHVILEAP 1969

RESULT 4

US-09-988-384B-33
 ; Sequence 33, Application US/0988384B
 ; Publication No. US20030073824A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, D.H.
 ; APPLICANT: Liu, H.
 ; APPLICANT: Xue, Y.
 ; APPLICANT: Zhao, L.
 ; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
 ; FILE REFERENCE: 600.536US1
 ; CURRENT APPLICATION NUMBER: US/09/988,384B
 ; CURRENT FILING DATE: 2001-11-19
 ; PRIOR APPLICATION NUMBER: PCT/US99/14398
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: US 09/105,537
 ; PRIOR FILING DATE: 1998-06-26
 ; NUMBER OF SEQ ID NOS: 53
 ; SEQ ID NO 33
 ; LENGTH: 3739
 ; TYPE: PRT
 ; ORGANISM: Streptomyces venezuelae
 US-09-988-384B-33

Query Match 71.9%; Score 1628.5; DB 9; Length 3739;
 Best Local Similarity 72.6%; Pred. No. 2.5e-119;
 Matches 312; Conservative 42; Mismatches 75; Indels 1; Gaps 1;

QY 1 EPIAIVGMACRPFPGVTSADDFMDLISSEBDALGCFPTDGMDDLTLYDPPDPHPCYCT 60
 DB 1540 EPAIVGMACRLPGVASPEDLWRLVAGSDAISGFPQDRGMVDEGLYDPPDPSGRTYC 1599
 QY 61 RNCGFLYDAGHPDAEPFGISPREALAMPQORLLLETAMETIEHAGINPTLHGTPTGVF 120
 DB 1600 RAGGFLDEAGFDDAFGISEPREALAMPQORLLLETAMETIEHAGIDPTSLGQOQGVF 1659
 QY 121 TGTNGDYALRVHNAQSTDFGALTGTAGSVISGRISYTFEGFEPAYSVDTACSSSLVAL 180
 DB 1660 AGTNGPYEPLNTAEDLEGVGTGNAASIMSGRVSYTLGLBPAVTVDTCSSSLVAL 1719
 QY 181 HLAQALRAGECSMALAGVTWSSPCAIFYEFSSRQGLADGHCKAFSAADGTGMEGV 240
 DB 1720 HLAVALRKESCCIALAGVTWSTPTTFVEFSRQGLAEDGRSKAFASADGFGPAEGV 1779
 QY 241 GMLIVERLSDAHNRGHRVLAIVRGSANODGASNGLTAPNGPSOQRYIRQALANAGSAG 300
 DB 1780 GMLIVERLSDARRNGHRVLAIVRGSANODGASNGLTAPNGPSOQRYIRRLADARLT7A 1839
 QY 301 DVDVAEAGTGTTLGDPTEAQAALLATYGGDRAGEGPMLGSVKSNVGHTOAAAGVAGV 360
 DB 1840 DVDVVEAHGCTGTRLDPIEAQAALLATYGGDRDTEQPLRLGSLKSNIGHTQAAGVSGITK 1899
 QY 361 MVMALRHGLLPRTLHVDPEPHVDMASAGAVQLTETVPMW-GGEGRLRAGVSSFGVSGT 419
 DB 1900 MVMAMRHGVLPKTLHVDPEPHVDMASAGAVQLTETVPMWPRKQGLRRAAVSSFGISGT 1959
 QY 420 NAHVILEAP 429
 DB 1960 NAHVILEAP 1969

RESULT 5

US-09-836-821-33
 ; Sequence 33, Application US/09836821
 ; Publication No. US20030087405A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, D.H.

APPLICANT: Liu, H.
 APPLICANT: Xue, Y.
 APPLICANT: Zhao, L.
 TITLE OF INVENTION: DNA encoding methymycin and pikromycin
 FILE REFERENCE: 600.438US1
 CURRENT APPLICATION NUMBER: US/09/836,821
 CURRENT FILING DATE: 2001-04-17
 PRIOR APPLICATION NUMBER: 09/105,537
 PRIOR FILING DATE: 1998-06-26
 NUMBER OF SEQ ID NOS: 43
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 33
 LENGTH: 3739
 TYPE: PRT
 ORGANISM: Streptomyces venezuelae
 US-09-836-821-33

Query Match 71.9%; Score 1628.5; DB 9; Length 3739;
 Best Local Similarity 72.6%; Pred. No. 2.5e-119;
 Matches 312; Conservative 42; Mismatches 75; Indels 1; Gaps 1;

QY 1 EPIAIVGMACRPFPGVTSADDFMDLISSEBDALGCFPTDGMDDLTLYDPPDPHPCYCT 60
 DB 1540 EPAIVGMACRLPGVASPEDLWRLVAGSDAISGFPQDRGMVDEGLYDPPDPSGRTYC 1599
 QY 61 RNCGFLYDAGHPDAEPFGISPREALAMPQORLLLETAMETIEHAGINPTLHGTPTGVF 120
 DB 1600 RAGGFLDEAGFDDAFGISEPREALAMPQORLLLETAMETIEHAGIDPTSLGQOQGVF 1659
 QY 121 TGTNGDYALRVHNAQSTDFGALTGTAGSVISGRISYTFEGFEPAYSVDTACSSSLVAL 180
 DB 1660 AGTNGPYEPLNTAEDLEGVGTGNAASIMSGRVSYTLGLBPAVTVDTCSSSLVAL 1719
 QY 181 HLAQALRAGECSMALAGVTWSSPCAIFYEFSSRQGLADGHCKAFSAADGTGMEGV 240
 DB 1720 HLAVALRKESCCIALAGVTWSTPTTFVEFSRQGLAEDGRSKAFASADGFGPAEGV 1779
 QY 241 GMLIVERLSDAHNRGHRVLAIVRGSANODGASNGLTAPNGPSOQRYIRQALANAGSAG 300
 DB 1780 GMLIVERLSDARRNGHRVLAIVRGSANODGASNGLTAPNGPSOQRYIRRLADARLT7A 1839
 QY 301 DVDVAEAGTGTTLGDPTEAQAALLATYGGDRAGEGPMLGSVKSNVGHTOAAAGVAGV 360
 DB 1840 DVDVVEAHGCTGTRLDPIEAQAALLATYGGDRDTEQPLRLGSLKSNIGHTQAAGVSGITK 1899
 QY 361 MVMALRHGLLPRTLHVDPEPHVDMASAGAVQLTETVPMW-GGEGRLRAGVSSFGVSGT 419
 DB 1900 MVMAMRHGVLPKTLHVDPEPHVDMASAGAVQLTETVPMWPRKQGLRRAAVSSFGISGT 1959
 QY 420 NAHVILEAP 429
 DB 1960 NAHVILEAP 1969

RESULT 6

US-09-793-708-2
 ; Sequence 2, Application US/09793708
 ; Publication No. US20030104537A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashley, Gary
 ; APPLICANT: BETHLACH, Melanie C.
 ; APPLICANT: BETLACH, Mary C.
 ; APPLICANT: MCDANIEL, Robert
 ; APPLICANT: TANG, Li
 ; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
 ; FILE REFERENCE: 30062202121
 ; CURRENT APPLICATION NUMBER: US/09/793,708
 ; CURRENT FILING DATE: 2001-02-22
 ; PRIOR APPLICATION NUMBER: US 09/657,440
 ; PRIOR FILING DATE: 2000-09-07
 ; PRIOR APPLICATION NUMBER: US 09/320,878
 ; PRIOR FILING DATE: 1999-05-27
 ; PRIOR APPLICATION NUMBER: US 09/141,908

PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: US 09/073,538
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: US 08/846,247
PRIOR FILING DATE: 1997-04-30
PRIOR APPLICATION NUMBER: US 60/134,990
PRIOR FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 2
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-793-708-2

Query Match 71.9%; Score 1628.5; DB 9; Length 3739;
Best Local Similarity 72.6%; Pred. No. 2.5e-119;
Matches 312; Conservative 42; Mismatches 75; Indels 1; Gaps 1;

QY 1 EPVAVGMACRPGVTSADDFMDLISSEODAIAGFPPTDRGMDLTLYPDPPHPTGCT 60
DB 1540 EPVAVGMACRLPGVASPEDLMLVAGGEDAISGFQDRGMVDEGLYDPPDASGRTYC 1599
QY 61 RAGGFYDAGHFDAPFFGISPREALAMPQORLLLETAMETIEHAGINPHTLHGTPTGVF 120
DB 1600 RAGGFLEAGEFDADFFGISPREALAMPQORLLLETSMWEAVDAGIDPTSLQGQGVGF 1659
QY 121 TGTNGDYALRVNNAAGSTGDFALITGAGSVISGRISTYTFGEPRAVSVDTACSSSLVAL 180
DB 1660 AGTNGPYEPLRLNTADLEBGVGTGNAASIMSGRVSYTLGLEBPAYVTVDACSSSLVAL 1719
QY 181 HLAQALRAGECSMALAGVTWSSPGAFFERSRORGLAADGHCKAFSAADGTGMGEV 240
DB 1720 HLAQALRKEGCGALAGVTWSTPTTFEFSRORGLAEDGRKAFSAADGTGMGEV 1779
QY 241 GMLIVERLSDARNRGRVLA VNRGS AVNODGASNGLTAPNPSQORVIRQALANAGISAG 300
DB 1780 GMLIVERLSDARNRGRVLA VNRGS AVNODGASNGLTAPNPSQORVIRQALANAGISAG 1839
QY 301 DVDAVEAHGCTTLDPIEAOALATTYGODRAGEPLMGSVSNVGHQTQAAAGVAVIK 360
DB 1840 DVDAVEAHGCTTLDPIEAOALATTYGODRAGEPLMGSVSNVGHQTQAAAGVAVIK 1899
QY 361 MWALRHGGLPRTLHVDEPSPHVMSAGAVQTLTETVPMW-GGEGRLRAGVSSFGVSGT 419
DB 1900 MWAMRHGGLPRTLHVDRPSQIDMSAGTYELLTEAMDMWRKQEGRLRAAVSSFGISGT 1959
QY 420 NAHVILEAP 429
DB 1960 NAHVILEAP 1969

RESULT 7
US-09-861-289-33
Sequence 33, Application US/09861289
Patent No. US20020110897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/861,289
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae

US-09-861-289-33

Query Match 71.9%; Score 1628.5; DB 10; Length 3739;
Best Local Similarity 72.6%; Pred. No. 2.5e-119;
Matches 312; Conservative 42; Mismatches 75; Indels 1; Gaps 1;

QY 1 EPVAVGMACRPGVTSADDFMDLISSEODAIAGFPPTDRGMDLTLYPDPPHPTGCT 60
DB 1540 EPVAVGMACRLPGVASPEDLMLVAGGEDAISGFQDRGMVDEGLYDPPDASGRTYC 1599
QY 61 RAGGFYDAGHFDAPFFGISPREALAMPQORLLLETAMETIEHAGINPHTLHGTPTGVF 120
DB 1600 RAGGFLEAGEFDADFFGISPREALAMPQORLLLETSMWEAVDAGIDPTSLQGQGVGF 1659
QY 121 TGTNGDYALRVNNAAGSTGDFALITGAGSVISGRISTYTFGEPRAVSVDTACSSSLVAL 180
DB 1660 AGTNGPYEPLRLNTADLEBGVGTGNAASIMSGRVSYTLGLEBPAYVTVDACSSSLVAL 1719
QY 181 HLAQALRAGECSMALAGVTWSSPGAFFERSRORGLAADGHCKAFSAADGTGMGEV 240
DB 1720 HLAQALRKEGCGALAGVTWSTPTTFEFSRORGLAEDGRKAFSAADGTGMGEV 1779
QY 241 GMLIVERLSDARNRGRVLA VNRGS AVNODGASNGLTAPNPSQORVIRQALANAGISAG 300
DB 1780 GMLIVERLSDARNRGRVLA VNRGS AVNODGASNGLTAPNPSQORVIRQALANAGISAG 1839
QY 301 DVDAVEAHGCTTLDPIEAOALATTYGODRAGEPLMGSVSNVGHQTQAAAGVAVIK 360
DB 1840 DVDAVEAHGCTTLDPIEAOALATTYGODRAGEPLMGSVSNVGHQTQAAAGVAVIK 1899
QY 361 MWALRHGGLPRTLHVDEPSPHVMSAGAVQTLTETVPMW-GGEGRLRAGVSSFGVSGT 419
DB 1900 MWAMRHGGLPRTLHVDRPSQIDMSAGTYELLTEAMDMWRKQEGRLRAAVSSFGISGT 1959
QY 420 NAHVILEAP 429
DB 1960 NAHVILEAP 1969

RESULT 8
US-09-860-846-6
Sequence 6, Application US/09860846
Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/860,846
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-860-846-6

Query Match 71.9%; Score 1628.5; DB 9; Length 11877;
Best Local Similarity 72.6%; Pred. No. 1.1e-118;
Matches 312; Conservative 42; Mismatches 75; Indels 1; Gaps 1;

QY 1 EPVAVGMACRPGVTSADDFMDLISSEODAIAGFPPTDRGMDLTLYPDPPHPTGCT 60
DB 6489 EPVAVGMACRLPGVASPEDLMLVAGGEDAISGFQDRGMVDEGLYDPPDASGRTYC 6548
QY 61 RAGGFYDAGHFDAPFFGISPREALAMPQORLLLETAMETIEHAGINPHTLHGTPTGVF 120
DB 6549 RAGGFLEAGEFDADFFGISPREALAMPQORLLLETSMWEAVDAGIDPTSLQGQGVGF 6608


```
RESULT 11
US-09-988-384B-6
; Sequence 6, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D. H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988, 384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 6
; LENGTH: 12199
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-6

Query Match      71.9%; Score 1628.5; DB 9; Length 12199;
Best Local Similarity 72.6%; Pred. No. 1,1e-118;
Matches 312; Conservative 42; Mismatches 75; Indels 1; Gaps 1;

QY 1 EPIAIVGMACRPPGVTASADPFWLISSEODAIIGFPDGRWDIDTLYPDPPHPTCYT 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 6811 EPVAIVGMACRLPGVAVSPEDLMRLVAGGDAISGFPDGRWDVEGYDPPDASGRTYC 6870
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 RANGFLVDAGHFDPAFFEGISPREALAMPQORLLLETAMETIEMAGINPTLHGTPTGVF 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 6871 RAGGFLEBAGFEDADFGISPREALAMPQORLLLETAMETIEMAGINPTLHGTPTGVF 6930
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 TGTNGODYALRVHNAAGSTDFALVTAGSVISGRISYTFGFBPAVSVDTACSSSLVAL 180
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 6931 AGTNGPHYEPRLRTAEDLEGVGTGNASTMSGRVSTLTGLEBPATVTDIACSSSLVAL 6990
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 HLAQALRAGECSMALAGVTWSSPGAFFEFSPORGLADGCKAFSAADGTGMEGV 240
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 6991 HLAQALRKEGCGALAGVTWSTPTTFVFSRORGLAEDGRKAPASADGFPAGEV 7050
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 GMLIVERLSDAHNRGHVLAIVRGSAVNODGASNGLTAPNGPSQORVIRALANAGISAG 300
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 7051 GMLIVERLSDARRNGHRLAVVRSANODGASNGLTAPNGPSQORVIRALADARLTTA 7110
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 DVDVAEAGTGTTGDPTEAQUALATYGGDRAGEGPMUGSVKSNVGHTOAAAGVAVIK 360
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 7111 DVDVVEAHGTGTRLDPIEAQALATYGGGRDTEQPLRLSLKNIGHTOAAAGVSGIIK 7170
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 MMALRHGGLPRTLHVDEPSPHVDWSAGAVOLLTETVPMP-GGEGRLRAGVSSFGVSGT 419
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 7171 MVQAMRHGGLPRTLHVDEPSDQIDWSAGAVELLTEADWPKQCGRLRAAVSSFGISGT 7230
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 420 NAHVILEEAP 429
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 7231 NAHVILEEAP 7240
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
US-09-793-708-1
; Sequence 1, Application US/09793708
; Publication No. US20030104597A1
; GENERAL INFORMATION:
; APPLICANT: Ashley, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: McDaniel, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002121
```

```

; CURRENT APPLICATION NUMBER: US/09/793, 708
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 09/657,440
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: US 09/141,908
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: US 09/073,538
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: US 08/846,247
; PRIOR FILING DATE: 1997-04-30
; PRIOR APPLICATION NUMBER: US 60/134,990
; PRIOR FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4551
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-793-708-1

Query Match      71.8%; Score 1625.5; DB 9; Length 4551;
Best Local Similarity 72.3%; Pred. No. 5,4e-119;
Matches 311; Conservative 41; Mismatches 77; Indels 1; Gaps 1;

QY 1 EPIAIVGMACRPPGVTASADPFWLISSEODAIIGFPDGRWDIDTLYPDPPHPTCYT 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1047 EPVAIVGMACRLPGVAVSPEDLMRLVAGGDAISGFPDGRWDVEGYDPPDASGRTYC 1106
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 RANGFLVDAGHFDPAFFEGISPREALAMPQORLLLETAMETIEMAGINPTLHGTPTGVF 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1107 RAGGFLEBAGFEDADFGISPREALAMPQORLLLETAMETIEMAGINPTLHGTPTGVF 1166
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 TGTNGODYALRVHNAAGSTDFALVTAGSVISGRISYTFGFBPAVSVDTACSSSLVAL 180
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1167 TGMTHEYGSLDGGEGLDGYLLTNTJASVMSGRVSYTLGLEBPALTVDIACSSSLVAL 1226
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 HLAQALRAGECSMALAGVTWSSPGAFFEFSPORGLADGCKAFSAADGTGMEGV 240
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1227 HLAQALRKEGVDMALAGVAVMPFPMFVFSRORGLADGCKAFSAADGTGMEGV 1286
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 GMLIVERLSDAHNRGHVLAIVRGSAVNODGASNGLTAPNGPSQORVIRALANAGISAG 300
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1287 GMLIVERLSDARRNGHRLAVVRSANODGASNGLTAPNGPSQORVIRALADARLTTA 1346
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 DVDVAEAGTGTTGDPTEAQUALATYGGDRAGEGPMUGSVKSNVGHTOAAAGVAVIK 360
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1347 DVDVVEAHGTGTRLDPIEAQALATYGGGRDTEQPLRLSLKNIGHTOAAAGVSGIIK 1406
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 MMALRHGGLPRTLHVDEPSPHVDWSAGAVOLLTETVPMP-GGEGRLRAGVSSFGVSGT 419
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1407 MVQAMRHGGLPRTLHVDEPSDQIDWSAGAVELLTEADWPKQCGRLRAAVSSFGISGT 1466
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 420 NAHVILEEAP 429
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1467 NAHVILEEAP 1476
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
US-09-860-846-31
; Sequence 31, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D. H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860, 846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
```

```

; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRF
; ORGANISM: Streptomyces venezuelae
US-09-860-846-31

```

```

Query Match      71.6%; Score 1622.5; DB 9; Length 4613;
Best Local Similarity 72.1%; Pred. No. 9.5e-119;
Matches 310; Conservative 42; Mismatches 77; Indels 1; Gaps 1;

```

```

QY 1 EPIAIVGACRPGCVTSADDFWDLISSEODAI GGFPTDRGMDLDTLYDPDPHPTCYT 60
D 1109 EPVAIVGACRLPGCVASPEDMLRLVAGGDAISEFPQDRGMDVGLYHPDPHPTSYV 1168
QY 61 RNGGFLYDAGHPDAEFFGISPREALAMPQORLLLETWETIEHAGINPHTLHGTPTGVF 120
D 1169 ROGGFIEHVAGFDAAFFGISPREALAMPQORLLLETWETIEHAGINPHTLHGTPTGVF 1228
QY 121 TGTNGODYALRVHNAAGSTGDFALTGTAGSVISGRISYTFGFEPAVSVPTACSSSLVAL 180
D 1229 TGANTHEYGSLRDGGGLDGYLLTGNTASVMSGRVSTTLGLEPALTVDPTACSSSLVAL 1288
QY 181 HLAQALRAGECSMALAGVTVMSSPGAFFVFSRQRLAADGHKAFSAADGTGMEGV 240
D 1289 HLAVALRKEGVMDALAGVAVMPTPGMFVFSRQRLAADGHKAFSAADGTGMEGV 1348
QY 241 GMLIVERLSDAHRNGHRLAVVNGSAVNOGASNGLTAPNGPSQORVIRQALANAGLSAG 300
D 1349 GVLIVERLSDARRNGHQLAVVNGSALNOGASNGLTAPNGPSQORVIRQALANAGLSAG 1408
QY 301 DVDAVEAHGCTTLDGPIEAOALLATYGORAGGPIMLGSVKNVNGHTQAAAGVAVIK 360
D 1409 DVDVEAHGCTRLGDIIEAOALLATYGQGRDDEQPLRLGSLKSNIGHTQAAAGVAVIK 1468
QY 361 MMVALRHGLPRTLHVDEPSPHVDMSAGAVOLLTETVPWP-GGEGRLRRAGVSSFGVSGT 419
D 1469 MVQAMRGLPKTLHVDEPSPQIDMSAGAVELTETVAVDMPEKQDGLRRRAVSSFGISGT 1528
QY 420 NAHVILEEAP 429
D 1529 NAHVILEEAP 1538

```

```

RESULT 14
US-09-988-384B-31
; Sequence 31, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D. H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRF
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-31

```

```

Query Match      71.6%; Score 1622.5; DB 9; Length 4613;
Best Local Similarity 72.1%; Pred. No. 9.5e-119;
Matches 310; Conservative 42; Mismatches 77; Indels 1; Gaps 1;

```

```

QY 1 EPIAIVGACRPGCVTSADDFWDLISSEODAI GGFPTDRGMDLDTLYDPDPHPTCYT 60
D 1109 EPVAIVGACRLPGCVASPEDMLRLVAGGDAISEFPQDRGMDVGLYHPDPHPTSYV 1168
QY 61 RNGGFLYDAGHPDAEFFGISPREALAMPQORLLLETWETIEHAGINPHTLHGTPTGVF 120
D 1169 ROGGFIEHVAGFDAAFFGISPREALAMPQORLLLETWETIEHAGINPHTLHGTPTGVF 1228
QY 121 TGTNGODYALRVHNAAGSTGDFALTGTAGSVISGRISYTFGFEPAVSVPTACSSSLVAL 180
D 1229 TGANTHEYGSLRDGGGLDGYLLTGNTASVMSGRVSTTLGLEPALTVDPTACSSSLVAL 1288
QY 181 HLAQALRAGECSMALAGVTVMSSPGAFFVFSRQRLAADGHKAFSAADGTGMEGV 240
D 1289 HLAVALRKEGVMDALAGVAVMPTPGMFVFSRQRLAADGHKAFSAADGTGMEGV 1348
QY 241 GMLIVERLSDAHRNGHRLAVVNGSAVNOGASNGLTAPNGPSQORVIRQALANAGLSAG 300
D 1349 GVLIVERLSDARRNGHQLAVVNGSALNOGASNGLTAPNGPSQORVIRQALANAGLSAG 1408
QY 301 DVDAVEAHGCTTLDGPIEAOALLATYGORAGGPIMLGSVKNVNGHTQAAAGVAVIK 360
D 1409 DVDVEAHGCTRLGDIIEAOALLATYGQGRDDEQPLRLGSLKSNIGHTQAAAGVAVIK 1468
QY 361 MMVALRHGLPRTLHVDEPSPHVDMSAGAVOLLTETVPWP-GGEGRLRRAGVSSFGVSGT 419
D 1469 MVQAMRGLPKTLHVDEPSPQIDMSAGAVELTETVAVDMPEKQDGLRRRAVSSFGISGT 1528
QY 420 NAHVILEEAP 429
D 1529 NAHVILEEAP 1538

```

```

RESULT 15
US-09-836-821-31
; Sequence 31, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D. H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: PRF
; ORGANISM: Streptomyces venezuelae
US-09-836-821-31

```

```

Query Match      71.6%; Score 1622.5; DB 9; Length 4613;
Best Local Similarity 72.1%; Pred. No. 9.5e-119;
Matches 310; Conservative 42; Mismatches 77; Indels 1; Gaps 1;
QY 1 EPIAIVGACRPGCVTSADDFWDLISSEODAI GGFPTDRGMDLDTLYDPDPHPTCYT 60
D 1109 EPVAIVGACRLPGCVASPEDMLRLVAGGDAISEFPQDRGMDVGLYHPDPHPTSYV 1168
QY 61 RNGGFLYDAGHPDAEFFGISPREALAMPQORLLLETWETIEHAGINPHTLHGTPTGVF 120
D 1169 ROGGFIEHVAGFDAAFFGISPREALAMPQORLLLETWETIEHAGINPHTLHGTPTGVF 1228
QY 121 TGTNGODYALRVHNAAGSTGDFALTGTAGSVISGRISYTFGFEPAVSVPTACSSSLVAL 180
D 1229 TGANTHEYGSLRDGGGLDGYLLTGNTASVMSGRVSTTLGLEPALTVDPTACSSSLVAL 1288

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 12:54:32 ; Search time 7.81547 Seconds
(without alignments)
1618.821 Million cell updates/sec

Title: US-09-914-286-3_COPY_2086_2515

Perfect score: 2265
Sequence: 1 EPIAIVGMACRPFQGVTSAD.....VSPFGVSGTNAHVILBAPA 430

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents AA.*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep.*
7: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1704	75.2	3170	4	US-09-036-987A-4
2	1704	75.2	3170	4	US-09-370-700-4
3	1672	73.8	4928	4	US-09-036-987A-5
4	1672	73.8	4928	4	US-09-370-700-5
5	1654	73.0	3519	4	US-09-428-517-4
6	1646.5	72.7	4150	4	US-09-428-517-2
7	1640	72.4	1996	2	US-08-804-227C-9
8	1640	72.4	1996	2	US-08-804-198-3
9	1640	72.4	5588	4	US-09-036-987A-6
10	1640	72.4	5588	4	US-09-370-700-6
11	1631.5	72.0	4472	2	US-08-804-227C-2
12	1628.5	71.9	3739	3	US-09-320-878-2
13	1628.5	71.9	3739	3	US-09-105-537-3
14	1628.5	71.9	11877	4	US-09-105-537-6
15	1625.5	71.8	4551	3	US-09-320-878-1
16	1624.5	71.7	4554	2	US-08-804-227C-14
17	1624.5	71.7	4550	2	US-08-804-227C-8
18	1624.5	71.7	4550	2	US-08-804-198-2
19	1622.5	71.6	2595	4	US-09-036-987A-2
20	1622.5	71.6	2595	4	US-09-370-700-2
21	1622.5	71.6	4613	4	US-09-105-537-31
22	1619.5	71.5	2152	4	US-09-036-987A-3
23	1619.5	71.5	2152	4	US-09-370-700-3
24	1609	71.0	3816	4	US-09-428-517-3
25	1597	70.5	3729	2	US-08-804-227C-4
26	1587.5	70.1	3724	2	US-08-804-227C-10
27	1587.5	70.1	3724	2	US-08-804-198-4

28	1576.5	69.6	1891	2	US-08-804-227C-12	Sequence 12, Appl
29	1576.5	69.6	1891	2	US-08-804-198-6	Sequence 6, Appl
30	1574.5	69.5	1841	2	US-08-804-227C-6	Sequence 6, Appl
31	1570	69.3	5215	4	US-09-105-537-2	Sequence 2, Appl
32	1561.5	68.9	1562	3	US-09-320-878-3	Sequence 3, Appl
33	1561.5	68.9	1562	3	US-09-105-537-35	Sequence 35, Appl
34	1544.5	68.2	3567	2	US-07-642-734C-4	Sequence 4, Appl
35	1544.5	68.2	3567	3	US-08-439-009A-2	Sequence 4, Appl
36	1543	68.1	1346	3	US-09-320-878-4	Sequence 4, Appl
37	1538	67.9	1346	4	US-09-105-537-37	Sequence 37, Appl
38	1535	67.8	1864	2	US-08-804-227C-3	Sequence 3, Appl
39	1515	66.9	1611	2	US-08-804-227C-5	Sequence 5, Appl
40	1495.5	66.0	3491	2	US-07-642-734C-2	Sequence 2, Appl
41	1495.5	66.0	3491	3	US-08-439-009A-2	Sequence 2, Appl
42	1448	63.9	1580	2	US-08-804-227C-11	Sequence 11, Appl
43	1448	63.9	1580	2	US-08-804-198-5	Sequence 5, Appl
44	1431.5	63.2	3170	2	US-07-642-734C-5	Sequence 5, Appl
45	1431.5	63.2	3170	3	US-08-439-009A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-036-987A-4
Sequence 4, Application US/09036987A
Patent No. 6143526
GENERAL INFORMATION:
APPLICANT: Balez, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Paul J.
APPLICANT: Turner, Van R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes for Spinosyn Insecticide
NUMBER OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036, 987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3170 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-4
Query Match 75.2%; Score 1704; DB 4; Length 3170;
Best Local Similarity 75.2%; Pred. No. 3, se-147;
Matches 323; Conservative 38; Mismatches 66; Indels 2; Gaps 1;

QY 1 EPIAIVGACRPGCVTSADDFMILISSEODAIIGFPTDRGMDLDTLYDPDPDHGTCTY 60
DB 34 EPIAIVMGCRYPGCVOPDEGMKLVASGDAIGFPPDRGMHLDLDPDPDGTCTY 93
QY 61 RRGFLYDAGHDEAFEGISPREALAMPQORLLLETMETIETHAGINPHIAGTPTGVF 120
DB 94 RHGGFLHDAGEFDAGFDISPREALAMPQORLLLETIETWETVESAGMDPRSLRGSRTGVF 153
QY 121 TGTNGQDYALAVHNAGOSTDGFALTGTGAGVISGRISTTFEGEPRAVSVDTACSSLYAL 180
DB 154 AGLMYEGYDTGAHRAEGVEGYLTGNAGSVASGVAAVAFEGEPAAVTVDTACSSLYAL 213
QY 181 HLAQALRABECSMALAGVTVMSSPGAFVEFSRORGLAADGCKAFSAADGTGMEGV 240
DB 214 HLAQOSLRQGBCDLALAGVTVMSTPERFVEFSRORGLAPOPCKRSPAAADGTGMEGA 273
QY 241 GMLVERLSAHRNGHRLAVVRSANVQDASNGLTAPNGPSQORVITROLANAGLSAG 300
DB 274 GLVILLERLSDARRNGHRLAVVRSANVQDASNGLTAPNGLAQERVIQVLTSAGLSAS 333
QY 301 DVDVAEAGTGTGDIPEAQLLATYGGDRAGEBPLMIGSVKSNVHTQAAAGVAVIK 360
DB 334 DVDVAEAGTGTGDIPEAQLLATYGGDRDRRPLMIGSVKSNIGHTQAAAGVAVIK 393
QY 361 MMALRHGLPRTLHVDESPHVDMSAGAVOLLTETVPMGEGRLRAGVSSFGVSGTN 420
DB 394 MMAMRHGELPRTLHVDEPNSHVDMSAGAVRLLTENIRMG--TGTTRAGVSSFGVSGTN 451
QY 421 AHVILEAP 429
DB 452 AHVILEHP 460

RESULT 2
US-09-370-700-4
Sequence 4, Application US/09370700
Patent No. 6274350
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patci J
APPLICANT: Turner, Jan R
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 4
LENGTH: 3170
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-4

Query Match 75.2%; Score 1704; DB 4; Length 3170;
Best Local Similarity 75.3%; Pred. No. 3.5e-147;
Matches 323; Conservative 38; Mismatches 66; Indels 2; Gaps 1;
QY 1 EPIAIVGACRPGCVTSADDFMILISSEODAIIGFPTDRGMDLDTLYDPDPDHGTCTY 60
DB 34 EPIAIVMGCRYPGCVOPDEGMKLVASGDAIGFPPDRGMHLDLDPDPDGTCTY 93
QY 61 RRGFLYDAGHDEAFEGISPREALAMPQORLLLETMETIETHAGINPHIAGTPTGVF 120
DB 94 RHGGFLHDAGEFDAGFDISPREALAMPQORLLLETIETWETVESAGMDPRSLRGSRTGVF 153
QY 121 TGTNGQDYALAVHNAGOSTDGFALTGTGAGVISGRISTTFEGEPRAVSVDTACSSLYAL 180

DB 154 AGLMYEGYDTGAHRAEGVEGYLTGNAGSVASGRVAVAFEGEPAAVTVDTACSSLYAL 213
QY 181 HLAQALRABECSMALAGVTVMSSPGAFVEFSRORGLAADGCKAFSAADGTGMEGV 240
DB 214 HLAQOSLRQGBCDLALAGVTVMSTPERFVEFSRORGLAPOPCKRSPAAADGTGMEGA 273
QY 241 GMLVERLSAHRNGHRLAVVRSANVQDASNGLTAPNGPSQORVITROLANAGLSAG 300
DB 274 GLVILLERLSDARRNGHRLAVVRSANVQDASNGLTAPNGLAQERVIQVLTSAGLSAS 333
QY 301 DVDVAEAGTGTGDIPEAQLLATYGGDRAGEBPLMIGSVKSNVHTQAAAGVAVIK 360
DB 334 DVDVAEAGTGTGDIPEAQLLATYGGDRDRRPLMIGSVKSNIGHTQAAAGVAVIK 393
QY 361 MMALRHGLPRTLHVDESPHVDMSAGAVOLLTETVPMGEGRLRAGVSSFGVSGTN 420
DB 394 MMAMRHGELPRTLHVDEPNSHVDMSAGAVRLLTENIRMG--TGTTRAGVSSFGVSGTN 451
QY 421 AHVILEAP 429
DB 452 AHVILEHP 460

RESULT 3
US-09-036-987A-5
Sequence 5, Application US/09036987A
Patent No. 6143526
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patci J.
APPLICANT: Turner, Jan R.
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4928 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-5

Query Match 73.8%; Score 1672; DB 4; Length 4928;
Best Local Similarity 75.2%; Pred. No. 5.8e-144;

Matches 322; Conservative 34; Mismatches 70; Indels 2; Gaps 2;

QY 3 IAIYGMACRPFPGVTSADDFWDLISSBODAIIGFPTDRGMDLTLYPDPPHPTGCTYRN 62
DB 1787 VAIYGMCRPFPGVTSPEELMRLVAGVDVAGFPDRGMDLALYDPDPRLGTSYCE 1846
QY 63 GGFLYDAGHPDAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLHGTPTGVTG 122
DB 1847 GGFLRDAAEFPDAMFGISPREALAMPQORLLLETAMETTERAGIDPFSLHGSRTGVAG 1906
QY 123 TNGODYALR-VHNAQSTIDGFPALGTAGSVISGRISYTFPREGPAVSVDTCACSSSLVALH 181
DB 1907 LMHYDYGARFTRPAEGFEGHGTGNAGSVLSGRVAVSPFEGPAVYVDTACSSSLVALH 1966
QY 182 LACQALRAGECSMALAGGVTVMSPGAUFVEFSRORGLADGHCAPSAADGTGMEGCVG 241
DB 1967 LAGCALRAGECEFPALAGGVTVMSPTTFVEFSRORGLAPDRCKSPFAAADGTGMEGAG 2026
QY 242 MLVERLSDAHRNGHRYLAVVRGSAVNODASNGLTAPNGPSQORVIRQALANAGLSAGD 301
DB 2027 LVLLERLSDAHRNGHEVLAVVRGSAVNODASNGLTAPNGPSQORVITQALTSAGLSVD 2086
QY 302 VDAVEAHGTGTTLDPIEAQALLATYGODRAGEBPLMGSVKNVGHTOAAAGVAVYIKM 361
DB 2087 VDAVEAHGTGTRLDPIEAQALLATYGRDRDPRPLMGSVKNIGHTOAAAGVAVYIKM 2146
QY 362 VMALRHGLPRTLHVDEPSPHVMSAGAVOLLTETVPMPCGEGRLRAGVSPFGVSGTNA 421
DB 2147 VMAMRQELPRTLHVDEPSAQVDSAGTVOLLTENTPMP-DSGRLRAGVSPFGISGTNA 2205
QY 422 HVILEEAP 429
DB 2206 HLILEEPP 2213

RESULT 4
US-09-370-700-5
Sequence 5, Application US/09370700
Patent No. 6274350
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madhuri, Krishnamurthy
APPLICANT: Treedway, Patti J
APPLICANT: Turner, Jan R
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Bioynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 4928
TYPE: PRT
ORGANISM: Saccharopolypora spinosa
US-09-370-700-5

Query Match 73.8%; Score 1672; DB 4; Length 4928;
Best Local Similarity 75.2%; Pred. No. 5.8e-144;
Matches 322; Conservative 34; Mismatches 70; Indels 2; Gaps 2;

QY 3 IAIYGMACRPFPGVTSADDFWDLISSBODAIIGFPTDRGMDLTLYPDPPHPTGCTYRN 62
DB 1787 VAIYGMCRPFPGVTSPEELMRLVAGVDVAGFPDRGMDLALYDPDPRLGTSYCE 1846
QY 63 GGFLYDAGHPDAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLHGTPTGVTG 122
DB 1847 GGFLRDAAEFPDAMFGISPREALAMPQORLLLETAMETTERAGIDPFSLHGSRTGVAG 1906

QY 123 TNGODYALR-VHNAQSTIDGFPALGTAGSVISGRISYTFPREGPAVSVDTCACSSSLVALH 181
DB 1907 LMHYDYGARFTRPAEGFEGHGTGNAGSVLSGRVAVSPFEGPAVYVDTACSSSLVALH 1966
QY 182 LACQALRAGECSMALAGGVTVMSPGAUFVEFSRORGLADGHCAPSAADGTGMEGCVG 241
DB 1967 LAGCALRAGECEFPALAGGVTVMSPTTFVEFSRORGLAPDRCKSPFAAADGTGMEGAG 2026
QY 242 MLVERLSDAHRNGHRYLAVVRGSAVNODASNGLTAPNGPSQORVIRQALANAGLSAGD 301
DB 2027 LVLLERLSDAHRNGHEVLAVVRGSAVNODASNGLTAPNGPSQORVITQALTSAGLSVD 2086
QY 302 VDAVEAHGTGTTLDPIEAQALLATYGODRAGEBPLMGSVKNVGHTOAAAGVAVYIKM 361
DB 2087 VDAVEAHGTGTRLDPIEAQALLATYGRDRDPRPLMGSVKNIGHTOAAAGVAVYIKM 2146
QY 362 VMALRHGLPRTLHVDEPSPHVMSAGAVOLLTETVPMPCGEGRLRAGVSPFGVSGTNA 421
DB 2147 VMAMRQELPRTLHVDEPSAQVDSAGTVOLLTENTPMP-DSGRLRAGVSPFGISGTNA 2205
QY 422 HVILEEAP 429
DB 2206 HLILEEPP 2213

RESULT 5
US-09-428-517-4
Sequence 4, Application US/09428517
Patent No. 6251636
GENERAL INFORMATION:
APPLICANT: Beclach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/428,517
CURRENT FILING DATE: 1999-10-28
EARLIER APPLICATION NUMBER: 60/120,254
EARLIER FILING DATE: 1999-02-16
EARLIER APPLICATION NUMBER: 60/106,100
EARLIER FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 3519
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-428-517-4

Query Match 73.0%; Score 1654; DB 4; Length 3519;
Best Local Similarity 73.3%; Pred. No. 1.6e-142;
Matches 315; Conservative 40; Mismatches 75; Indels 0; Gaps 0;

QY 1 EPIAIYGMACRPFPGVTSADDFWDLISSBODAIIGFPTDRGMDLTLYPDPPHPTGCTYRN 60
DB 1687 EAIATAMSCRFPGGJDSPEDLWFLAEGDAVAGLEPDRGMDLALYHDPDPENPGTTYV 1746
QY 61 RINGPLYDAGHPDAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLHGTPTGVF 120
DB 1747 REGAFRYDAQFPAAGFPGISPREALAMPQORLLLETAMETTERAGIDPFSLHGSRTGVAG 1806
QY 121 TNGODYALR-VHNAQSTIDGFPALGTAGSVISGRISYTFPREGPAVSVDTCACSSSLVAL 180
DB 1807 IGAGHGYGDPDRGRAPESVAGVILYGTASAVLSGRISYTFGLEGPAVYVDTACSSSLVAL 1866
QY 181 HLAQALRAGECSMALAGGVTVMSPGAUFVEFSRORGLADGHCAPSAADGTGMEGCVG 240
DB 1867 HLAVQALRAGECSLAAGVAVVMSDPAFVEFSRORGLADGHCAPSAADGTGMEGCVG 1926

Qy 241 GMLIVERLSDAHNRHGRVLA VVRSASVNOGASNGLTAPNGPSQORVIRQALANAGLSAG 300
Db 1927 STLLERLSDARLRGHVLA VVRSASVNOGASNGLTAPNGPSQORVIRQALANAGLSAG 1986
Qy 301 DVDAVEAHGCTGTRLDPIEAQALLATTGQDPAGSGPLMLGSVKNVNGHTQAAAGVAGYIK 360
Db 1987 DVDAVEAHGCTGTRLDPIEAQALLATTGQDPAGSGPWLGSVKNIGHTQAAAGVAGYMK 2046
Qy 361 MVMALRHGLPRTLHVDPSPHVDMWSAGAVOLLTETVPMGPGEGLRIRAGVSSFGVSGTN 420
Db 2047 MVLALGRGVKPLTHVDEPSPHVDMWSAGAVALLTEERPEERLRIRAGISARGVSGTN 2106
Qy 421 AHVILEEAPA 430
Db 2107 AHVILEEAPA 2116

RESULT 6

US-09-428-517-2
Sequence 2, Application US/09428517
Patent No. 6251636
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/428.517
CURRENT FILING DATE: 1999-10-28
EARLIER APPLICATION NUMBER: 60/120.254
EARLIER FILING DATE: 1999-02-16
EARLIER APPLICATION NUMBER: 60/106.100
EARLIER FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4150
TYPE: PR1
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-428-517-2

Query Match 72.7%; Score 1646.5; DB 4; Length 4150;
Best Local Similarity .72.8%; Pred. No. 9.9e-142;
Matches 313; Conservative 41; Mismatches 75; Indels 1; Gaps 1;

Qy 1 EPIAIVGACRFPFGVTSADDFMDLISSEODAIAGFPTRDGMDDTLVDPDPHPTCYT 60
Db 2546 EPIAIVGACRFPFGVTSADDFMDLISSEODAIAGFPTRDGMDDTLVDPDPHPTCYT 2605
Qy 61 RNGGFLVDAGHFDPAFFGDISPREALAMPDQORLLLETAMETIEHAGINPHTLHPTGVF 120
Db 2606 REGAFLHAGHFDPAFFGDISPREALAMPDQORLLLETAMETIEHAGINPHTLHPTGVF 2665
Qy 121 TGTNGQVYALRVNAGSGTDGFLATGTAAGTISGRISTTFGBEPNAYSVDTCAGSSLYAL 180
Db 2666 TGTNGQVYALRVNAGSGTDGFLATGTAAGTISGRISTTFGBEPNAYSVDTCAGSSLYAL 2725
Qy 181 HLAQALRAGECSMALAGVTWSSPGAFFVFSRORGLAADGHCKAFSAADGTGMEGV 240
Db 2726 HLAQALRAGECSMALAGVTWSSPGAFFVFSRORGLAADGHCKAFSAADGTGMEGV 2785
Qy 241 GMLIVERLSDAHNRHGRVLA VVRSASVNOGASNGLTAPNGPSQORVIRQALANAGLSAG 300
Db 2786 GMLIVERLSDAHNRHGRVLA VVRSASVNOGASNGLTAPNGPSQORVIRQALANAGLSAG 2845
Qy 301 DVDAVEAHGCTGTRLDPIEAQALLATTGQDPAGSGPLMLGSVKNVNGHTQAAAGVAGYIK 360
Db 2846 DVDAVEAHGCTGTRLDPIEAQALLATTGQDPAGSGPWLGSVKNIGHTQAAAGVAGYMK 2905

Qy 361 MVMALRHGLPRTLHVDPSPHVDMWSAGAVOLLTETVPMGPGEGLRIRAGVSSFGVSGTN 420
Db 2906 MVLALGRGVKPLTHVDEPSPHVDMWSAGAVALLTEERPEERLRIRAGISARGVSGTN 2964
Qy 421 AHVILEEAPA 430
Db 2965 AHVILEEAPA 2974

RESULT 7

US-08-804-227C-9
Sequence 9, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Roestek, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM Compatible
SOFTWARE: ASCII (DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804.227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1996 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-227C-9

Query Match 72.4%; Score 1640; DB 2; Length 1996;
Best Local Similarity 71.9%; Pred. No. 1.4e-141;
Matches 309; Conservative 46; Mismatches 75; Indels 0; Gaps 0;

Qy 1 EPIAIVGACRFPFGVTSADDFMDLISSEODAIAGFPTRDGMDDTLVDPDPHPTCYT 60
Db 35 EPIAIVGACRFPFGVTSADDFMDLISSEODAIAGFPTRDGMDDTLVDPDPHPTCYT 94
Qy 61 RNGGFLVDAGHFDPAFFGDISPREALAMPDQORLLLETAMETIEHAGINPHTLHPTGVF 120
Db 95 REGAFLHAGHFDPAFFGDISPREALAMPDQORLLLETAMETIEHAGINPHTLHPTGVF 154
Qy 121 TGTNGQVYALRVNAGSGTDGFLATGTAAGTISGRISTTFGBEPNAYSVDTCAGSSLYAL 180
Db 155 TGTNGQVYALRVNAGSGTDGFLATGTAAGTISGRISTTFGBEPNAYSVDTCAGSSLYAL 214
Qy 181 HLAQALRAGECSMALAGVTWSSPGAFFVFSRORGLAADGHCKAFSAADGTGMEGV 240
Db 215 HLAQALRAGECSMALAGVTWSSPGAFFVFSRORGLAADGHCKAFSAADGTGMEGV 274
Qy 241 GMLIVERLSDAHNRHGRVLA VVRSASVNOGASNGLTAPNGPSQORVIRQALANAGLSAG 300
Db 275 GMLIVERLSDAHNRHGRVLA VVRSASVNOGASNGLTAPNGPSQORVIRQALANAGLSAG 334

```

Qy 301 DVDAAVEAHGTGTTLDDPFIIEAOLLLATYGGDRAGEGLMGYSKSNVGHQAAAGYAVTX 360
Db 335 DVDAAVEAHGTGTPLEDPIBAGLLLATYSGERGQGPMLGSLKSNIGHAQAAGYGVYIK 394
Qy 361 MWMLRHGLLPRTTHVDEBSPHYDWSAGAVOLLTETVPWPGEGRLRRAGVSSFGVSGTN 420
Db 395 VVQAMRHGSLPRTLHVDAPOSKEVNAAGAVELLTETRSPRIVERRRAAVAFVSGTN 454
Qy 421 AHVILEEAPA 430
Db 455 AHVILEEAPA 464

RESULT 8
US-08-804-198-3
/ Sequence 3, Application US/08804198
/ Patent No. 5945320
/ GENERAL INFORMATION:
/ APPLICANT: Burgett, Stanley G.
/ APPLICANT: Kufstoss, Stuart A.
/ APPLICANT: Ric, Nagaraja R.
/ APPLICANT: Richardson, Mark A.
/ APPLICANT: Rostock, Paul R., Jr.
/ TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: PAUL R. CANTRELL, 1138
/ STREET: LILLY CORPORATE CENTER
/ CITY: INDIANAPOLIS
/ STATE: IN
/ COUNTRY: USA
/ ZIP: 46285
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: Macintosh
/ OPERATING SYSTEM: Macintosh 7.0
/ SOFTWARE: Microsoft Word 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/804,198
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CANTRELL, PAUL R.
/ REGISTRATION NUMBER: 36,470
/ REFERENCE/DOCKET NUMBER: P9113
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 317-276-3885
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1996 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: peptide
US-08-804-198-3

Query Match 72.4%; Score 1640; DB 2; Length 1996;
Best Local Similarity 71.9%; Pred. No. 1.4e-141;
Matches 309; Conservative 46; Mismatches 75; Indels 0; Gaps 0;

Qy 1 EPPIAVGMAACPRGCVTSADDFWDLISSEQDAIGCFPTDRGMDLDTLYDDPDHFGCTYT 60
Db 35 EPPIAVGMAACPRGCVTSADDFWDLISSEQDAIGCFPTDRGMDLDTLYDDPDHFGCTYT 94
Qy 61 RRGGLFLVYAGHDAEPFGISPREALAMPQOGLLETAMETETEHAGINPHLTGPTGVF 120
Db 95 RRGGLFLVYAGHDAEPFGISPREALAMPQOGLLETAMETETEHAGINPHLTGPTGVF 154
Qy 121 TGTNGQDYALRVHNAQGSITGDFALVTGTAGSVISGRISYTFEGEPGAVSVDTACSSSLVAL 180
Db 155 TGVWVDYDGSRDSAPPEYEGVLTVNGSAGSINSGRVAVALGLGEPALVTVDTCSSSLVAL 214
Qy 181 HLACQALAAEGESMALAGCVTMTSPGAFVETSRORGLAANDCHCAFAALADGTGMBEV 240

```

Db 215 HLA*V08:1R:RGECDDLLAAGCVTMA*PTVLVEFSPRGIIAADGRCKAFAEGADGTAMAEV 274
 Qy 241 GMLIVLRSLSDAHNRNHRVLAAYRGSAVNODASNGLTAPNPGSPQORVYIRQALNANAGLSAG 300
 Db 275 GVLVLRSLSDAHNRNHRVLAAYRGSAVNODASNGLTAPNPGSPQORVYIRQLADAGLTAPA 334
 Qy 301 DVDVAEAGGTGTTLDGPIEAQALLATTYQODRAGEGPIMLGSKVSNVGHNTQAAAGVAVIK 360
 Db 335 DVDVAEAGHTGTPLDPIEAQALLATTYSEKGGGPIMLGSKINIHQAAPAAAGVGVIK 394
 Qy 361 MVALAEHGLPRTLHYDEBSPHYDWSAGAVOLLTETVPWPCGEGEQLRRAGVSSFFVGSTN 420
 Db 395 VVQAMHAGSLPRTLHYDAPSSKVEMASGAVELLTETSPMRPRVRRVRAAIVSAFVGSGTN 454
 Qy 421 AHVYLEEAPA 430
 Db 455 AHVYLEEAPA 464

```

      RESULT 9
      US-09-036-987A-6
      Sequence 6, Application US/09036987A
      Patent No. 6143526
      GENERAL INFORMATION:
      APPLICANT: Baltz, Richard H.
      APPLICANT: Broughton, Mary C.
      APPLICANT: Crawford, Kathryn P.
      APPLICANT: Madduri, Krishnamurthy
      APPLICANT: Merlo, Donald J.
      APPLICANT: Treadway, Patli J.
      APPLICANT: Turner, Jan R.
      TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
      NUMBER OF INVENTION: Production
      NUMBER OF SEQUENCES: 39
      CORRESPONDENCE ADDRESSES:
      ADDRESSEE: Dow Agrosciences LLC Patent Department
      STREET: 9330 Zionsville Road
      CITY: Indianapolis
      STATE: Indiana
      COUNTRY: USA
      ZIP: 46268
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.30
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/036,987A
      FILING DATE: 09-MAR-1998
      CLASSIFICATION: 435
      ATTORNEY/AGENT INFORMATION:
      NAME: Stuart, Donald R
      REGISTRATION NUMBER: 28,479
      REFERENCE/DOCKET NUMBER: 50,608
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (317)337-4816
      TELEFAX: (317)337-4847
      INFORMATION FOR SEQ ID NO: 6:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 5588 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-09-036-987A-6

```

Db 1782 DPVIVGMACRFGPGVSSPEELMRLVAGGLDAVAEPDDRGMDQAGLFPDDPDLGTSYV 1841
QY 61 RANGELVDAGHDAEPFGISPREALAMPQORLLLETAMETIEHAGINPHLTGTPGVF 120
Db 1842 CEGGFLLDAAEFDAEPFGISPREALAMPQORLLLETAMETIEHAGINPHLTGTPGVF 1901
QY 121 TGTNGODYALR-VNNAAGOSTDGFALGTAGSVISGRISYTPGFEPAVSVDTCSSSLVA 179
Db 1902 AGLMHHYGAARFIRAPBEGFVIGNSAGVSGRVAVSGFPGPAVTDVTCSSSLVA 1961
QY 180 LHLAQLRAGECSMALAGVTVWSSPGAFVEFSRORGLAADGCKAFSAADGTGMEG 239
Db 1962 LHLAQLRAGECSMALAGVTVWSSPGAFVEFSRORGLAADGCKAFSAADGTGMEG 2021
QY 240 VGMILVERLSDAHRNGHRLVAVRGSANODGASNGLTAPNGSPQORVITQALASAGLSV 299
Db 2022 AGVLVERLSDAHRNGHRLVAVRGSANODGASNGLTAPNGSPQORVITQALASAGLSV 2081
QY 300 GDVDAVEAHGTGTLGDPTEAQLLATYGGDRAGEPLMGSVKSNVGTQAAAGVAVI 359
Db 2082 SDVDAVEAHGTGTLGDPTEAQLLATYGGDRSDRPLMGSVKSNIGHTQAAAGVAVI 2141
QY 360 KMWALRHGLPRTLHVDESPHYDWSAGAVOLLTEVPWPGGRLRRAGVSSFGVSGT 419
Db 2142 KMWAMRHGQLPATLHVDEPTEVDWSAGDVOLLTEVPMP-GNSHPRRVGVSSFGISGT 2200
QY 420 NAHVILEEA 428
Db 2201 NAHVILEEA 2209

RESULT 10

US-09-370-6
; Sequence 6, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:-
; APPLICANT: Ballez, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn. Ver. 2.0
; SEQ ID NO 6
; LENGTH: 5588
; TYPE: PRT
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-6

Query Match 72.4%; Score 1640; DB 4; Length 5588;
Best Local Similarity 73.2%; Pred. No. 6e-141;
Matches 314; Conservative 40; Mismatches 73; Indels 2; Gaps 2;

QY 1 EPIAIVMACRFGPGVTSADDFMDLISSEDAIGGFPTRGMDLDTLYDPDPDHPTCYT 60
Db 1782 DPVIVGMACRFGPGVSSPEELMRLVAGGLDAVAEPDDRGMDQAGLFPDDPDLGTSYV 1841
QY 61 RANGFLVDAGHDAEPFGISPREALAMPQORLLLETAMETIEHAGINPHLTGTPGVF 120
Db 1842 CEGGFLLDAAEFDAEPFGISPREALAMPQORLLLETAMETIEHAGINPHLTGTPGVF 1901
QY 121 TGTNGODYALR-VNNAAGOSTDGFALGTAGSVISGRISYTPGFEPAVSVDTCSSSLVA 179
Db 1902 AGLMHHYGAARFIRAPBEGFVIGNSAGVSGRVAVSGFPGPAVTDVTCSSSLVA 1961

QY 180 LHLAQLRAGECSMALAGVTVWSSPGAFVEFSRORGLAADGCKAFSAADGTGMEG 239
Db 1962 LHLAQLRAGECSMALAGVTVWSSPGAFVEFSRORGLAADGCKAFSAADGTGMEG 2021
QY 240 VGMILVERLSDAHRNGHRLVAVRGSANODGASNGLTAPNGSPQORVITQALASAGLSV 299
Db 2022 AGVLVERLSDAHRNGHRLVAVRGSANODGASNGLTAPNGSPQORVITQALASAGLSV 2081
QY 300 GDVDAVEAHGTGTLGDPTEAQLLATYGGDRAGEPLMGSVKSNVGTQAAAGVAVI 359
Db 2082 SDVDAVEAHGTGTLGDPTEAQLLATYGGDRSDRPLMGSVKSNIGHTQAAAGVAVI 2141
QY 360 KMWALRHGLPRTLHVDESPHYDWSAGAVOLLTEVPWPGGRLRRAGVSSFGVSGT 419
Db 2142 KMWAMRHGQLPATLHVDEPTEVDWSAGDVOLLTEVPMP-GNSHPRRVGVSSFGISGT 2200
QY 420 NAHVILEEA 428
Db 2201 NAHVILEEA 2209

RESULT 11

US-08-804-227C-2
; Sequence 2, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:-
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kuhstoes, Stuart A.
; APPLICANT: Roeseck, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:-
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:-
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: ASCII (DOS) Text only
; CURRENT APPLICATION DATA:-
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:-
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:-
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:-
; LENGTH: 4472 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-804-227C-2

Query Match 72.0%; Score 1631.5; DB 2; Length 4472;
Best Local Similarity 71.9%; Pred. No. 2.6e-140;
Matches 309; Conservative 46; Mismatches 74; Indels 1; Gaps 1;

QY 1 EPIAIVMACRFGPGVTSADDFMDLISSEDAIGGFPTRGMDLDTLYDPDPDHPTCYT 60
Db 2576 DPVIVGMACRFGPGVTSADDFMDLISSEDAIGGFPTRGMDLDTLYDPDPDHPTCYT 2635
QY 61 RANGFLVDAGHDAEPFGISPREALAMPQORLLLETAMETIEHAGINPHLTGTPGVF 120
Db 2636 REGGFLLDAAEFDAEPFGISPREALAMPQORLLLETAMETIEHAGINPHLTGTPGVF 2695

Qy	121	TGTNGODVATLVHNAGSGSTDEGALVTGTAYSVSGI	SVTFPGREGSAVVDTRACSSVLAL	180
Db	2696	VGLTHQETASRIHEAPEEYEGILLTKGKASVVS	SGIISTTLGUEGSLSDTRACSSVLAL	2755
Qy	181	HLACQALPAGBECSMALAGVTVMS	SPGAFVEFSRQRLAADGHCKAFSAADGTGMEGV	240
Db	2756	HNAQALPAGBECDDMALAGVTVMA	RGLFVEFSRQRLAADGRCFAFDGADGTAMEGA	2815
Qy	241	GMLVERLSDAHRNCHRVLA	VVRGSAVNQDGSNGLTAPNGSPQQRVIRQALANAGLSAG	300
Db	2816	GVLIVERLSDARRLGH	PVTLAVVCGSAVNQDGSNGLTAPSGSPQQRVIRQALANARLTVA	2875
Qy	301	DVDAVEANGSTGTTACDPIEAQ	ALLATTYQODRAGBEPMLGSKYSNVGHTQAAAGVATYK	360
Db	2876	DVDVVEANGSTGRLCDPIEAQ	ALLLTGYGRDRACFPWLGSKLSKSI	GHAQAAAGVATYK 2935
Qy	361	MVMALRHGLLPRTLHVDESP	PHVMSAGAVOLLTETVMPGSEGLRBRAGSSFGVSGTN	420
Db	2936	MVLARVYQMLPRTLHVDESP	RRHVMSAGAVRLLTETAREMPGVN	KPRRAVASARFVSGTN 2994
Qy	421	AHVILEEAPA	430	
Db	2995	AHLITAEAPA	3004	

```

RESULT 12
US-09-330-878-2
; Sequence 2, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320, 878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141, 908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073, 538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846, 247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119, 139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100, 880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087, 080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 3739
; TYPR: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-320-878-2

Query Match      71.9%; Score 1628.5; DB 3; Length 3739;
Beat Local Similarity 72.6%; Pred. No. 3.8e-140;
Matches 312; Conservative 42; Mismatches 75; Indels 1; Gaps 1;

Cy      1  EPPIAVGNACRRPGGCVTASDADFMDLISEQDAIGCFPFIDRGMDLDTLYDPPDHGTCYT 60
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1540  BPVALVGNACRIPGGVAPBBDLMLRVAGGEPAISGFPDDRGMDEGLYDPPDASGRTYC 15999

Cy      61  RNGGFLVYAGHFDADFPGGISPREALAMPQOORLLLETMETETIEHGINPHILHGTPTGVF 120
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      1600  RAGGFLDAGGERDADFPGISPREALAMPQOORLLLETETMEAVEDAGIDPTSLQCGQGVCF 16558

Cy      121  TGNGQDYALRVHNAAGSTDGFALTGTAGSVISGRISTYFPGEGEPAVSVDTACSSSLVAL 180

```

Db	1660	AGTNGPHYEPLNTAEDLEGVYCTGAASIMGRSVYTLGLEPNAVYDTJACSSSLVAL	1719
Qy	181	HLACQALRAGECSMALAGVYTWSSPGAFYEFSRQRLAADGHCIAFSAADGTGWEGV	240
Db	1720	HLAQALRKCECGIALAGVYTWSTPTTFVEFESRQRLAADGSKAFMAADGFGPAGV	1779
Qy	241	GMLIVELLSAHRNGHVLAIVVGSANODGASNGCLAPMGPSQOQRYIRPLALDARLTTA	300
Db	1780	GMLIVELLSQRRRNGHVLAIVVGSANODGASNGCLAPMGPSQOQRYIRPLALDARLTTA	1839
Qy	301	DVDAVEAHGTGTLGPPIEAQALLATYGGDRAGEGPLMLGSVSNVGHTOAAAGVAVIK	360
Db	1840	DVDVVEAHGTGTRGPIIEAQALLATYGGDRDTEQPLRGLSLSKNIGHTQAAAGVSVLIK	1899
Qy	361	MMALRLRGLLPRLIHDVDESPHYDMSAGAVQLTETVPMW-GGEGRLRRGVSSFGVSGT	419
Db	1900	MYQAMRGVGLPKLHVDPRDQIDWSAGVELLTLEAMDWERQKGGGLRRVAVSFGISGT	1959
Qy	420	NAHYILEEAP	429
Db	1960	NAHYILEEAP	1969

```

RESULT 13
US-09-105-537-33
: Sequence 33, Application US/09105537A
: Patent No. 6265202
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600.438US1
: CURRENT APPLICATION NUMBER: US/09/105,537A
: CURRENT FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 33
: LENGTH: 3739
: TYPE: RPT
: ORGANISM: Streptomyces venezuelae
US-09-105-537-33

```

Query Match	71.94%	Score 1628.5	DB 4	Length 3739
Best Local Similarity	72.64%	Pred. No. 3.8e-140		
Matches 312	Conservative 42	Mismatches 75	Indels 1	Gaps 1
Qy	1	EPPIAVGMACRFPGCVTSADPFMDLLISSEODAIAGFPEFDRGMDLDTLYDDPDPHFGCYT	60	
Db	1540	EPVAIVGMACRLPGVAVSPEDLWRLVAGSGDAISGFPQDGMVDEGLYDDPDPAASGTYC	1599	
Qy	61	RNGSFLVDAGHPDAEPFGISPREALAMPQORLLLETATWETIEHAGINPTLHGTPTGVF	120	
Db	1600	RAGGFLDEAGEPADDFPGISPREALAMPQORLLLETSMVAEVDAGIDPTSLGQGVGF	1659	
Qy	121	TGTNGQDYALRVHNAAGSTDGPAALTGNAISGRISYTFEGFGPANSVDTACSSSLVAL	180	
Db	1660	AGTNGPHYEPRLRTADELGEVGTGPAASIMSGRVSYTLGLSGPATVDTACSSSLVAL	1719	
Qy	181	HLACQALRAGECSMALAGVTWASSPGAFEFESRORRLADGHCKAFSAADAGTGMCEGV	240	
Db	1720	HLAVQALRKKECGALAGVTWVSTPTTFEFSRQRLADGHSKAPAAADAGGPAEGV	1779	
Qy	241	GMLIVERLSPAHNRGRVLAVTRGSANODGASNGLTAPNGPSQQRVIRQALNAGLSAG	300	
Db	1780	GMLIVERLSPARRNGHVLVAVRGSANOCGASNGLTAPNGPSQQRVIRBALDLARLTTA	1839	
Qy	301	DVAVERHNGGTLTGDPTEAQAALLATYGDRBAGEPLWGSVSNVGHDTAAAGVGVIK	360	
Db	1840	DVDVVEAHNGGTRGDPIEAQAALLATYGGQRDTQPPRLGSLNSNIGHTPAAAGVSITIK	1899	

Oy	361	MWMLRHGLLPRRTLTDEPSPHVDWSACAVOLLTETVMP-CGEGRRLRAGGSSPSVGST	419
Db	1900	MVGMARHGVLKXTHLVDRPSDQIDWSAGTVELTEAMDMPPRGEGTLRAAIVSSIGISGT	1955
Oy	420	NAHTVLEEAR	429
Db	1960	NAHTVLEEAR	1969

```

RESULT 14
US-09-105-537-6
: Sequence 6, Application US/09105537A
: Patent No. 6265302
: GENERAL INFORMATION:
: APPLICANT: Sherman, D. H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: TITLE OF INVENTION: DNA encoding myhmcy1
: FILE REFERENCE: 600.438051
: CURRENT APPLICATION NUMBER: US/09/105,537A
: CURRENT FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 11877
: TYPE: CDS
: ORGANISM: Streptomyces venezuelae
US-09-105-537-6

```

	Query Match Similarity	71.9%;	Score 1628.5;	DB 4;	Length 11877;				
	Best Local Similarity	72.6%;	Pred. No. 2e-119;						
Matches	312;	Conservative	42;	Mismatches	75;	Indels	1;	Gaps	1.
QY	1	EPPIAVGNACPRPGSVTSADPFWDLISSEQDAIGCFPTDRGNDLDTTYIDPPDPHFQCTYT	60						
Db	6489	EPVALVGNACRLPGGVASPEDLMRWLVAGGBDAISGFPODRGMDVGLVDPDPDASGRITYC	6548						
QY	61	RNGGFLVYAGHFDAEPFGISPREALAMPQOQRLILETAMETLEHAGINPHTLHAGPTGVF	120						
Db	6549	RAGGFLDADGERFDADFGISPREALAMPQOQRLILETSMENAEVDAGIDPTSLQSGQGVF	6600						
QY	121	TGTNGQDYALRVHNAQSGTDEGAFALTGTNAGSVISGRISTTFGEFGPAVSVDITACSSLYAL	180						
Db	6609	AGTNGPHEPPLIRNIAEDLEGVVGNGNAASIMSGVSYYTLGLEGPVVYDTACSSLYAL	6668						
QY	181	HLAQQALRAGECSMYLACGVTMTSSPGAFFVERSRGRGLAADCHCAFSAAADGTMGRCV	240						
Db	6669	HLAVQALRKSGCGLLACGVTMTSPFTTFVERSRGRGLAEDORSKAFPAASADGFPARCV	6722						
QY	241	GMLLVERLSDAHRNHRVLAIVVRGSAVNQDGAISNGLTAPNGBSQQRVIRQALANAGLSAG	300						
Db	6729	GMLLVERISDARRNHRVLAIVVRGSAVNQDGAISNGLTAPNGPSQQRVIRRALADARLTYYA	6788						
QY	301	DVDVAEHAAGTGTTGSDPIEAQALLATYQODRAGEPLMGYSKSVNWHGTQAAAGVAYIX	360						
Db	6789	DVDVAEHAAGTGTRLSDPIEAQALLATYQGGKRDTEPLRIGSKISNIGHQQAAGVSGIIX	6844						
QY	361	MYMALRHGLLPRTLTHVDESPHVDMSAGAVOLLITTEWMP-OGEGRLRRAGVSPFGVSGT	419						
Db	6849	MYQAMRHGVLPTTLTHVDRPDQIDWSAGVELLITRAMDMPRKQEGGLRRAAVSFGISGT	6908						
QY	420	NAHVILIEEAP 429							
Db	6909	NAHVILIEEAP 6918							

RESULT 15
US-09-320-878-1
; Sequence 1, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary

```

? APPLICANT: BETLACH, Melanie C.
? APPLICANT: BETLACH, Mary C.
? APPLICANT: MCDANIEL, Robert
? APPLICANT: TANG, Li
? TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
? FILE REFERENCE: 300622002120
? CURRENT APPLICATION NUMBER: US/09/320,879A
? CURRENT FILING DATE: 1999-05-27
? EARLIER APPLICATION NUMBER: CIP OF 09/141,908
? EARLIER FILING DATE: 1998-08-28
? EARLIER APPLICATION NUMBER: CIP OF 09/073,558
? EARLIER FILING DATE: 1998-05-06
? EARLIER APPLICATION NUMBER: CIP OF 08/846,247
? EARLIER FILING DATE: 1997-04-30
? EARLIER APPLICATION NUMBER: 60/119,139
? EARLIER FILING DATE: 1999-02-08
? EARLIER APPLICATION NUMBER: 60/100,880
? EARLIER FILING DATE: 1998-09-22
? EARLIER APPLICATION NUMBER: 60/087,080
? EARLIER FILING DATE: 1998-05-28
? NUMBER OF SEQ ID NOS: 34
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO: 1
? LENGTH: 4551
? TYPE: PR1
? ORGANISM: Streptomyces venezuelae
JS-09-320-878-1

```

	Query Match	Similarity	71.8%	Score	1625.5	DA 3	Length	4551
	Best Local	Similarity	72.3%	Pred.	No.	9.6e-140		
	Matches	311	Conservative	41	Mismatches	77	Indels	1
								Gaps 1
Qy	1	EPVAIVMACRPGCGVTSADDFMDLISSEDDAIGGFPTDGBMDLDTLYDDPDHPGTCY	60					
Db	1047	EPVAIVMACRLPGGVASPEDLWRLVVGGGDAISEFQDQGWVBGLYHPDPHPGTSYV	1106					
Qy	61	RNGGFLVDAGHPDDEFPGISPREALAMDPOORLLTFAMETIHBAGINPHLLGTPGVF	120					
Db	1107	RGGFPIENVAGFDPAFPGISPREALAMDPOORLLTFETSMNAVEDAGIDLPFSLGRQGVF	1166					
Qy	121	TGNGGVYALRVHNAVGSTDGAFITGAGSVISGRISYTFGFEGPAVSVDTACSSSLVAL	180					
Db	1167	TGAMTHFEGSLRPGCGEGLDGYLLTGMTASVMSGRVSYTLGLGEPALTVPVTCSSSLVAL	1226					
Qy	181	HLACQALRAPECSMALAGVTWVSSPCAFAFEFSRQRGLADGHCKAASAAADGQWMEGV	240					
Db	1227	HLAVALARKGEVDMLALGVAIVMPTPEMFAFEFSRQRGLADGSKAPAAADGTSMEGV	1286					
Qy	241	GMLLVERLSDAHRNGHVLAIVRGSAAVNQDASNGTLTPNGPSQOQRYIROLANAGLSAG	300					
Db	1287	GVLIVERLSDARRRNGHVLAIVRGSAAVNQDASNGTLTPNGPSQOQRYIRRALDARLTLS	1346					
Qy	301	DVDAVEAHNGTTLGDPTEAQAALLATGQDRAEGPLMTGVSXSVNCHTQAAGVAVIK	360					
Db	1347	DVDAVEAHNGTTLGDPTEAQAALLATGQGRDDEQPLRLSLSLNSNIHTQAAGVSVIK	1406					
Qy	361	MYMALRRGLLPTRLHAVDESPPHVDMSAGAVQLLTETVPWP-GEGRILRAAGVSSFGVSGT	419					
Db	1407	MYQMRGRLLPKTLHAVDESPDQIDMSAGAVEILLTEAVDMEPEKQDGLRRAAVSFGISGT	1466					
Qy	420	NAHYITLEAP 429						
Db	1467	NAHYITLEAP 1476						

Search completed: June 17, 2003, 13:09:29
Job time : 8.81548 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 12:50:57 ; Search time 21.0646 Seconds
(without alignments)
2720.102 Million cell updates/sec

Title: US-09-914-286-3_COPY_2086_2515
Perfect score: 2265
Sequence: 1 EPIAIVGMACRPFPGVTSAD.....VSSFGVSGTNHVLLEAPA 430

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: /SID92/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID92/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID92/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID92/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID92/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID92/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID92/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID92/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID92/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID92/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID92/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID92/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID92/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID92/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID92/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID92/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID92/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID92/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID92/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID92/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID92/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID92/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID92/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2265	100.0	3972	21	AA823749
2	2265	100.0	3972	22	AA655264
3	2265	100.0	3972	22	AA655268
4	2042	90.2	6239	21	AA823750
5	2042	90.2	6239	22	AA655265
6	1968.5	86.9	5532	21	AA823752
7	1968.5	86.9	5532	22	AA655267
8	1744.5	77.0	4881	21	AA823751
9	1744.5	77.0	4881	22	AA655266
10	1704	75.2	3170	20	AAV39299

11	1704	75.2	3170	22	AA870967	S. spinosa protein
12	1672	73.8	4924	22	AA870968	S. spinosa protein
13	1672	73.8	4928	20	AAV39300	Spn a polyketide
14	1656	73.1	11096	22	AA810129	Streptomyces nours
15	1654	73.0	3519	21	AAV92709	S. antibioticus 8
16	1646.5	72.7	4150	21	AAV92707	S. antibioticus 8
17	1645.5	72.6	6797	22	AA831558	Pimaricin biosynth
18	1640	72.4	1996	18	AA822607	Platenolide synthase
19	1640	72.4	1568	18	AAV39301	Spn a polyketide
20	1640	72.4	5588	20	AA870969	S. spinosa protein
21	1640	72.4	5588	22	AA870969	S. spinosa protein
22	1636	72.2	5069	19	AAV52846	A. mediterranei r1
23	1631.5	72.0	4472	18	AAV22601	Tylosone synthase
24	1628.5	71.9	3739	21	AA818638	Amino acid sequenc
25	1628.5	71.9	3739	21	AAV77193	S. venezuelae macr
26	1628.5	71.9	3739	21	AAV77201	S. venezuelae pik
27	1628.5	71.9	3739	21	AAV77202	Nardanolide synthase
28	1628.5	71.9	12199	21	AAV77180	S. venezuelae pik
29	1625.5	71.8	4551	21	AA818637	Amino acid sequenc
30	1625.5	71.8	4551	21	AA818637	Nardanolide synthase
31	1624.5	71.7	4545	18	AAW22611	Hybrid smg/cy1g O
32	1624.5	71.7	4550	18	AAW23716	Platenolide synthase
33	1624.5	71.7	4550	18	AAW22606	Platenolide synthase
34	1622.5	71.6	2535	20	AAV39297	Spn a polyketide
35	1622.5	71.6	2535	22	AA870965	S. spinosa protein
36	1622.5	71.6	4613	21	AAV77192	S. venezuelae macr
37	1622.5	71.6	4613	21	AAV77200	S. venezuelae pik
38	1619.5	71.5	2152	20	AAV39298	Spn a polyketide
39	1619.5	71.5	2152	22	AA870966	S. spinosa protein
40	1614.5	71.3	7068	22	AA810142	Streptomyces nours
41	1614.5	71.3	9472	22	AA810144	Streptomyces nours
42	1611	71.1	4572	19	AAV52845	A. mediterranei r1
43	1609	71.0	3816	21	AAV92708	S. antibioticus 8
44	1605	70.9	888	23	AAV76935	Hybrid polyketide
45	1597	70.5	3729	18	AAW22603	Tylosone synthase

ALIGNMENTS

RESULT 1
AAB23749 standard; Protein; 3972 AA.
ID AAB23749;
AC AAB23749;
DT 10-JAN-2001 (first entry)
XX
DE S. avermectilis avermectin aglycon synthase protein SEQ ID NO:3.
XX
XX Streptomyces avermectilis; avermectin aglycon synthase; biosynthesis;
KM multifunctional enzyme; polyketide; avermectin; veterinary drug;
KM agrochemical.
XX
XX Streptomyces avermectilis.
OS
XX
XX WO2000050605-A1.
PD
XX 31-AUG-2000.
XX
XX 23-FEB-2000; 2000MO-JP01041.
PF
XX 24-FEB-1999; 99JP-0046961.
PR
XX (KITA) KITASATO INST.
XX
XX Omura S, Ikeda H;
PI
XX WPI: 2000-565458/52.
DR N-PSDB; AAA92301.
XX
XX Avermectin aglycone synthase DNA and proteins encoded by all or part of
PT it for the production of avermectin and its derivatives for drug and

PT agrochemical use -
XX
PS Claim 32; Page 203-223; 314pp; Japanese.
XX
CC The present sequence represents an avermectin aglycon synthase protein.
CC Also described are: (1) polypeptides encoded by all or part of the DNA;
CC (2) expression vectors containing the DNA; (3) host cells transformed by
CC the vectors; (4) preparation of the polypeptides by culture of the
CC transformants; (5) preparation of avermectin aglycon or its derivatives
CC by culture of transformed avermectin-producing microorganisms; and (6)
CC oligonucleotides of 5-60 bases in length containing sense or antisense
CC sequences from the avermectin aglycon synthase DNA. The enzymes are
CC useful for the production of modified forms of avermectin and of the
CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
CC and agrochemicals.
XX
SQ Sequence 3972 AA;
Query Match 100.0%; Score 2265; DB 21; Length 3972;
Best Local Similarity 100.0%; Pred. No. 1.1e-192;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPIAIVGMACRFGGVTGTSADDFMDLISSEODAIIGFPTDGMDDLTLYPDPPHGTCTY 60
DB 2086 EPIAIVGMACRFGGVTGTSADDFMDLISSEODAIIGFPTDGMDDLTLYPDPPHGTCTY 2145
QY 61 RNSGFVLDAGHFDFAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLHGTPTGVF 120
DB 2146 RNSGFVLDAGHFDFAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLHGTPTGVF 2205
QY 121 TGTNGODYALRVNNAOSTDGFALTGTAGSVISGRISYTFEGEPGPAVSVDTCSSSLVAL 180
DB 2206 TGTNGODYALRVNNAOSTDGFALTGTAGSVISGRISYTFEGEPGPAVSVDTCSSSLVAL 2265
QY 181 HLAQALRAGECSMALAGVTVMSPGAFVFEFSRQGLADGCKAFSAADGTGMBGV 240
DB 2266 HLAQALRAGECSMALAGVTVMSPGAFVFEFSRQGLADGCKAFSAADGTGMBGV 2325
QY 241 GMLIVERLSDAHNRGHRVLA VVRGSAVNODGASNGLTAPNGSQORVITROALNAGISAG 300
DB 2326 GMLIVERLSDAHNRGHRVLA VVRGSAVNODGASNGLTAPNGSQORVITROALNAGISAG 2385
QY 301 DVDVAEHAHGTGTTGDPTEAQLLATYGGDRAGEGPIMLGSVKSNGHTQAAAGVAVIK 360
DB 2386 DVDVAEHAHGTGTTGDPTEAQLLATYGGDRAGEGPIMLGSVKSNGHTQAAAGVAVIK 2445
QY 361 MVMALRHGILPRTLHVDESPHYDMSAGAVOLLTETVPMGEGRLRRAGVSSFGVSGTN 420
DB 2446 MVMALRHGILPRTLHVDESPHYDMSAGAVOLLTETVPMGEGRLRRAGVSSFGVSGTN 2505
QY 421 AHVILEEAPA 430
DB 2506 AHVILEEAPA 2515
RESULT 2
ID AAG65264 standard; Protein; 3972 AA.
XX
AC AAG65264;
XX
DT 04-DEC-2001 (first entry)
XX
DE Streptomycetes avermilitilis protein SEQ ID NO: 4.
XX
KM Avermectin aglycone synthase; AAS; avermectin derivative;
KM drug production; veterinary drug; pesticide.
XX
OS Streptomycetes avermilitilis.
XX
PN WO200162939-A1.
XX
PD 30-AUG-2001.

XX
PF 23-FEB-2001; 2001WO-JP01381.
XX
PR 24-FEB-2000; 2000JP-0047405.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PA (KITA) KITASATO INST.
XX
PI Endo H, Yamauchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
XX
DR WPI; 2001-582053/65.
XX
DR N-PSDB; AAH9277.
XX
PT New modified avermectin aglycone synthase derived from Streptomycetes
PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
PT drugs and pesticides -
XX
PS Claim 4; Page 167-180; 257pp; Japanese.
XX
CC The present invention relates to the production of modified derivatives
CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
CC avermilitilis. The activity of an acyl carrier protein (ACP),
CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC suppressed. The process can be used in the production of drugs, veterinary
CC drugs and pesticides. The present sequence is an S. avermilitilis protein.
XX
SQ Sequence 3972 AA;
Query Match 100.0%; Score 2265; DB 22; Length 3972;
Best Local Similarity 100.0%; Pred. No. 1.1e-192;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPIAIVGMACRFGGVTGTSADDFMDLISSEODAIIGFPTDGMDDLTLYPDPPHGTCTY 60
DB 2086 EPIAIVGMACRFGGVTGTSADDFMDLISSEODAIIGFPTDGMDDLTLYPDPPHGTCTY 2145
QY 61 RNSGFVLDAGHFDFAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLHGTPTGVF 120
DB 2146 RNSGFVLDAGHFDFAEFPGISPREALAMPQORLLLETAMETIEHAGINPHTLHGTPTGVF 2205
QY 121 TGTNGODYALRVNNAOSTDGFALTGTAGSVISGRISYTFEGEPGPAVSVDTCSSSLVAL 180
DB 2206 TGTNGODYALRVNNAOSTDGFALTGTAGSVISGRISYTFEGEPGPAVSVDTCSSSLVAL 2265
QY 181 HLAQALRAGECSMALAGVTVMSPGAFVFEFSRQGLADGCKAFSAADGTGMBGV 240
DB 2266 HLAQALRAGECSMALAGVTVMSPGAFVFEFSRQGLADGCKAFSAADGTGMBGV 2325
QY 241 GMLIVERLSDAHNRGHRVLA VVRGSAVNODGASNGLTAPNGSQORVITROALNAGISAG 300
DB 2326 GMLIVERLSDAHNRGHRVLA VVRGSAVNODGASNGLTAPNGSQORVITROALNAGISAG 2385
QY 301 DVDVAEHAHGTGTTGDPTEAQLLATYGGDRAGEGPIMLGSVKSNGHTQAAAGVAVIK 360
DB 2386 DVDVAEHAHGTGTTGDPTEAQLLATYGGDRAGEGPIMLGSVKSNGHTQAAAGVAVIK 2445
QY 361 MVMALRHGILPRTLHVDESPHYDMSAGAVOLLTETVPMGEGRLRRAGVSSFGVSGTN 420
DB 2446 MVMALRHGILPRTLHVDESPHYDMSAGAVOLLTETVPMGEGRLRRAGVSSFGVSGTN 2505
QY 421 AHVILEEAPA 430
DB 2506 AHVILEEAPA 2515
RESULT 3
ID AAG65268 standard; Protein; 3972 AA.
XX
AC AAG65268;
XX

DT 04-DEC-2001 (first entry)
 XX Streptomyces avermectin protein derivative SEQ ID NO: 8.
 XX
 XX
 KM Avermectin aglycone synthase; AAS; avermectin derivative;
 KM drug production; veterinary drug; pesticide.
 XX
 OS Synthetic.
 XX
 PN WO200162939-A1.
 XX
 PD 30-AUG-2001.
 XX
 PF 23-FEB-2001; 2001WO-JP01381.
 XX
 PR 24-FEB-2000; 2000JP-0047405.
 XX
 PA (KYOWA) KYOWA HAKKO KOCYO KK.
 PA (KITA) KITASATO INST.
 XX
 PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
 XX
 DR WPI; 2001-582053/65.
 DR N-PSDB; AAH79279.
 XX
 PT New modified avermectin aglycone synthase derived from Streptomyces
 PT avermectin used in production of 22,23-dihydroavermectin B1a used in
 PT drug and pesticides -
 XX
 PS Claim 5; Page 235-248; 257pp; Japanese.
 XX
 CC The present invention relates to the production of modified derivatives
 CC of avermectin aglycone synthase (AAS) derived from Streptomyces
 CC avermectin. The activity of an acyl carrier protein (ACP)
 CC beta-ketolactone carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketolactone carrier protein reductase (KR), dehydratase (DH), enoyl
 CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is derived from an S.
 CC avermectin protein.
 XX
 XX
 SQ Sequence 3972 AA;
 Query Match 100.0%; Score 2265; DB 22; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 1,1e-192;
 Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPIAIVGACRPFPGVTSADDFWDLISSEODAGSPTRDGMWLDLTYDPDPHPTGCT 60
 DB 2086 EPIAIVGACRPFPGVTSADDFWDLISSEODAGSPTRDGMWLDLTYDPDPHPTGCT 2145
 QY 61 RINGFLYDAGHFDPAEFPGISPREALAMPQORLLLETAETIEHAGINPHTLHGTPTGVF 120
 DB 2146 RINGFLYDAGHFDPAEFPGISPREALAMPQORLLLETAETIEHAGINPHTLHGTPTGVF 2205
 QY 121 TGTNGODYALRVHNAQSGTDFGALTTGAGSVISGRISYTFGFEPAVSVDTACSSSLVAL 180
 DB 2206 TGTNGODYALRVHNAQSGTDFGALTTGAGSVISGRISYTFGFEPAVSVDTACSSSLVAL 2265
 QY 181 HLACQALRAGECSMALAGVTVMSPPGAPVEFSRQGLADGHCAPSAADGTGMEGV 240
 DB 2266 HLACQALRAGECSMALAGVTVMSPPGAPVEFSRQGLADGHCAPSAADGTGMEGV 2325
 QY 241 GMLLVERLSDAHRNGHVLAVVRGSAVNOGASNGLTRAPNGPSQORVIRQALNAGISAG 300
 DB 2326 GMLLVERLSDAHRNGHVLAVVRGSAVNOGASNGLTRAPNGPSQORVIRQALNAGISAG 2385
 QY 301 DVDVAEAGTCTTGDPIEAQALLATYGGDRAGGPIMLGSKVNGHTQAAAGVAVIK 360
 DB 2386 DVDVAEAGTCTTGDPIEAQALLATYGGDRAGGPIMLGSKVNGHTQAAAGVAVIK 2445
 QY 361 MVMALRHGLLPTLHVDESPHVDMSAGAVQLLTETVPMWPGEGRLRAGVSSFGVSGTN 420
 |||

DB 2446 MVMALRHGLLPTLHVDESPHVDMSAGAVQLLTETVPMWPGEGRLRAGVSSFGVSGTN 2505
 QY 421 AHVILEEAPR 430
 |||
 DB 2506 AHVILEEAPR 2515
 RESULT 4
 ID AAB23750
 AA AAB23750 strand; Protein; 6239 AA.
 AC AAB23750;
 XX
 DT 10-JAN-2001 (first entry)
 XX
 DE S. avermectin avermectin aglycone synthase protein SEQ ID NO:4.
 XX
 KM Streptomyces avermectin; avermectin aglycone synthase; biosynthesis;
 KM multifunctional enzyme; polyketide; avermectin; veterinary drug;
 KM agrochemical.
 XX
 OS Streptomyces avermectin.
 XX
 PN WO200050605-A1.
 XX
 PD 31-AUG-2000.
 XX
 PF 23-FEB-2000; 2000WO-JP01041.
 XX
 PR 24-FEB-1999; 99JP-0046961.
 XX
 PA (KITA) KITASATO INST.
 XX
 PI Omura S, Ikeda H;
 XX
 DR WPI; 2000-565458/52.
 DR N-PSDB; AAA92301.
 XX
 PT Avermectin aglycone synthase DNA and protein encoded by all or part of
 PT it for the production of avermectin and its derivatives for drug and
 PT agrochemical use -
 XX
 PS Claim 32; Page 223-254; 314pp; Japanese.
 XX
 CC The present sequence represents an avermectin aglycone synthase protein.
 CC Also described are: (1) polypeptides encoded by all or part of the DNA;
 CC (2) expression vectors containing the DNA; (3) host cells transformed by
 CC the vectors; (4) preparation of the polypeptides by culture of the
 CC transformants; (5) preparation of avermectin aglycone or its derivatives
 CC by culture of transformed avermectin-producing microorganisms; and (6)
 CC oligonucleotides of 5-60 bases in length containing sense or antisense
 CC sequences from the avermectin aglycone synthase DNA. The enzymes are
 CC useful for the production of modified forms of avermectin and of the
 CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
 CC and agrochemicals.
 XX
 XX
 SQ Sequence 6239 AA;
 Query Match 90.2%; Score 2042; DB 21; Length 6239;
 Best Local Similarity 89.5%; Pred. No. 1,1e-172;
 Matches 385; Conservative 14; Mismatches 31; Indels 0; Gaps 0;
 QY 1 EPIAIVGACRPFPGVTSADDFWDLISSEODAGSPTRDGMWLDLTYDPDPHPTGCT 60
 DB 2680 EPIAIVGACRPFPGVTSADDFWDLISSEODAGSPTRDGMWLDLTYDPDPHPTGCT 2739
 QY 61 RINGFLYDAGHFDPAEFPGISPREALAMPQORLLLETAETIEHAGINPHTLHGTPTGVF 120
 DB 2740 RINGFLYDAGHFDPAEFPGISPREALAMPQORLLLETAETIEHAGINPHTLHGTPTGVF 2739
 QY 121 TGTNGODYALRVHNAQSGTDFGALTTGAGSVISGRISYTFGFEPAVSVDTACSSSLVAL 180
 DB 2800 TGTNGODYALRVHNAQSGTDFGALTTGAGSVISGRISYTFGFEPAVSVDTACSSSLVAL 2859
 |||

```

QY 181 HLACQALRAGECSMALAGVTVMSSPGAFFVEFSRORGLAADGHCKAFSAADGTGMEGV 240
DB 2660 HLACQALRAGECSMALAGVTVMSSPGAFFVEFSRORGLAADGHCKAFSAADGTGMEGV 2919
QY 241 GMLVERLSDAHRNGHRLAVVRSANVODGASNGLTAPNGPSQORVIRQALANAGLSAG 300
DB 2920 GMLVERLSDAHRNGHRLAVVRSANVODGASNGLTAPNGPSQORVIRQALANADLTTPA 2979
QY 301 DVDAVEAHGCTTIGDPIEAQALLATYGDRPAGSGPLMGSVKSNGHTQAAAGVAVIK 360
DB 2980 DVDAVEAHGCTTIGDPIEAQALLATYGDRPAGSGPLMGSVKSNGHTQAAAGVAVIK 3039
QY 361 MVMALRHGLLPTLHVDEPSPHYDMSAGAVOLLTETVPMPCGEGRLRRAGVSFGVSGTN 420
DB 3040 MVMALRHGLLPTLHVDEPSPHYDMSAGAVOLLTETVPMPCGEGRLRRAGVSFGVSGTN 3099
QY 421 AHVILEBAPA 430
DB 3100 AHVILEBAPA 3109

RESULT 5
AAG65265 standard; Protein: 6239 AA.
ID AAG65265
XX
AC AAG65265;
XX
DT 04-DEC-2001 (first entry)
XX
DE Streptomycetes avermiltillis protein SEQ ID NO: 5.
XX
KM Avermectin aglycone synthase; AAs: avermectin derivative;
  drug production; veterinary drug; pesticide.
XX
OS Streptomycetes avermiltillis.
XX
PN WO200162939-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-JP01381.
XX
PR 24-FEB-2000; 2000JP-0047405.
XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
  (KITA ) KITASATO INST.
XX
PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
  WPI; 2001-582053/65.
DR N-PSDB; AAH79277.
XX
PT New modified avermectin aglycone synthase derived from Streptomycetes
  avermiltillis used in production of 22,23-dihydroavermectin B1a used in
  drugs and pesticides -
XX
PS Claim 4; Page 180-201; 257pp; Japanese.
XX
XX The present invention relates to the production of modified derivatives
  of avermectin aglycone synthase (AAS) derived from Streptomycetes
  avermiltillis. The activity of an acyl carrier protein (ACP),
  CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
  CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
  CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
  CC suppressed. The process can be used in the production of drugs, veterinary
  CC drugs and pesticides. The present sequence is an S. avermiltillis protein.
XX
SQ Sequence 6239 AA;
  Query Match 90.2%; Score 2042; DB 22; Length 6239;
  Best Local Similarity 89.5%; Pred. No. 1.8e-172;
  Matches 385; Conservative 14; Mismatches 31; Indels 0; Gaps 0;

```

```

QY 1 EPIAIVGMACRPPGVTSSADDFMDLISSEODATGFPTRGMDLDTLYDPPDPHPTCYT 60
DB 2680 EPIAIVGMACRPPGVTSSADDFMDLISSEODATGFPTRGMDLDTLYDPPDPHPTCYT 2739
QY 61 RNCGFLYDAGHPDAEFFGISPREALAMPQORLLLETAMETIBHAGINPHTLHGTPTGVF 120
DB 2740 RNCGFLYDAGHPDAEFFGISPREALAMPQORLLLETAMETIBHAGINPHTLHGTPTGVF 2799
QY 121 TGTNGQDYALRNHNAQSGTDGFPALTGTAGSVISGRISYTFPGFGRVAVSDTACSSSLVAL 180
DB 2800 TGTNGQDYALRNHNAQSGTDGFPALTGTAGSVISGRISYTFPGFGRVAVSDTACSSSLVAL 2859
QY 181 HLACQALRAGECSMALAGVTVMSSPGAFFVEFSRORGLAADGHCKAFSAADGTGMEGV 240
DB 2860 HLACQALRAGECSMALAGVTVMSSPGAFFVEFSRORGLAADGHCKAFSAADGTGMEGV 2919
QY 241 GMLVERLSDAHRNGHRLAVVRSANVODGASNGLTAPNGPSQORVIRQALANAGLSAG 300
DB 2920 GMLVERLSDAHRNGHRLAVVRSANVODGASNGLTAPNGPSQORVIRQALANADLTTPA 2979
QY 301 DVDAVEAHGCTTIGDPIEAQALLATYGDRPAGSGPLMGSVKSNGHTQAAAGVAVIK 360
DB 2980 DVDAVEAHGCTTIGDPIEAQALLATYGDRPAGSGPLMGSVKSNGHTQAAAGVAVIK 3039
QY 361 MVMALRHGLLPTLHVDEPSPHYDMSAGAVOLLTETVPMPCGEGRLRRAGVSFGVSGTN 420
DB 3040 MVMALRHGLLPTLHVDEPSPHYDMSAGAVOLLTETVPMPCGEGRLRRAGVSFGVSGTN 3099
QY 421 AHVILEBAPA 430
DB 3100 AHVILEBAPA 3109

RESULT 6
AAB23752 standard; Protein: 5532 AA.
ID AAB23752
XX
AC AAB23752;
XX
DT 10-JAN-2001 (first entry)
XX
DE S. avermiltillis avermectin aglycon synthase protein SEQ ID NO:6.
XX
KM Streptomycetes avermiltillis; avermectin aglycon synthase; biosynthesis;
  multifunctional enzyme; polyketide; avermectin; veterinary drug;
  agrochemical.
XX
OS Streptomycetes avermiltillis.
XX
PN WO200050605-A1.
XX
PD 31-AUG-2000.
XX
PF 23-FEB-2000; 2000WO-JP01041.
XX
PR 24-FEB-1999; 99JP-0046961.
XX
PA (KITA ) KITASATO INST.
XX
PI Omura S, Ikeda H;
  WPI; 2000-565458/52.
DR N-PSDB; AAA92302.
XX
PT Avermectin aglycone synthase DNA and proteins encoded by all or part of
  it for the production of avermectin and its derivatives for drug and
  agrochemical use -
XX
PS Claim 32; Page 278-305; 314pp; Japanese.
XX
XX The present sequence represents an avermectin aglycon synthase protein.
  Also described are: (1) polypeptides encoded by all or part of the DNA;

```


XX agrochemical.
 XX
 OS Streptomyces avermectilis.
 XX
 PN WO200050605-A1.
 XX
 PD 31-AUG-2000.
 XX
 PF 23-FEB-2000; 2000WO-JP01041.
 XX
 PR 24-FEB-1999; 99JP-0046961.
 XX
 PA (KITA) KITASATO INST.
 XX
 PI Omura S, Ikeda H;
 DR WPI; 2000-565458/52.
 XX
 N-PSDB; AAA92302.
 XX
 PT Avermectin aglycone synthase DNA and proteins encoded by all or part of
 PT it for the production of avermectin and its derivatives for drug and
 PT agrochemical use _
 XX
 PS Claim 32; Page 254-278; 314pp; Japanese.
 XX
 CC The present sequence represents an avermectin aglycon synthase protein.
 CC Also described are: (1) polypeptides encoded by all or part of the DNA;
 CC (2) expression vectors containing the DNA; (3) host cells transformed b
 CC the vectors; (4) preparation of the polypeptides by culture of the
 CC transformants; (5) preparation of avermectin aglycon or its derivatives
 CC by culture of transformed avermectin-producing microorganisms; and (6)
 CC oligonucleotides of 5-60 bases in length containing sense or antisense
 CC sequences from the avermectin aglycon synthase DNA. The enzymes are
 CC useful for the production of modified forms of avermectin and of the
 CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
 CC and agrochemicals.
 XX
 SO Sequence 4881 AA;

Query Match	77.0%: Score 1744.5; DB 21; Length 4881.
Best Local Similarity	76.7%: Pred. No. 5.3e-146;
Matches 329; Conservative	40; Mismatches 59; Indels 1; Gaps 1;
QY 1	EP1AIVGNACFPFGGVTSADDFWDLISSEODAIIGGFPDPRGMDLDTLTPDPDPDHGTCYCT 60
Db 2636	EP1AIVGNACFPFGGVSAEDPFWELIAGRDVAGFPPVDRGMDVEAFDPDEPGAGTST 26395
QY 61	RNGEFLVAGHPDAEFEGGISPREALAMPQORLLLETWETLEHAGINPHTLGTPTGVF 120
Db 2696	RCGGFLQGAAFEDAGFFGISPREALAMPQORLMEVSEWELERAGIDPATLHGTTGVF 27555
QY 121	TGTNGDDVALRVHNAAGOSTDGFALGTGATYSISGRISTYFGEGGAVSVDPACSSLTVAL 180
Db 2756	AGVSGDDVALRLRKTQDHEGALTGVSNVSNGSLSTTFGEFGPAVTVDPACSSLTVAL 28155
QY 181	HLACQALRAGECSMNLAGVTVMSSPGAFAVEFSRQGLIADGHCKAFSAADGTGWGEGV 240
Db 2816	HLACQALRSGECSLTALAGVTVMSTPGAFAVEFSRQGLSPQRCAYAGSADGVGMAGV 28755
QY 241	GMLVVERISDAHRNGHRLVAVYRGSANVQDASNGLTLPNGFSQQRVTRQALNANGLSAG 300
Db 2876	GVLVVERISEAERRHRVLAAYRGSANVQDASNGLTLPNGFSQQRVTRQALACGLSIVA 29355
QY 301	DVDVAEAGSTGTTGCDPIEAQALATLTYGQDRAEGEPLMLGSKSVNGHTQAAAGVAYIX 360
Db 2936	DVDVVEGGGTGTTGCDPIEAQDLATLTYQGSGGERPWLGSKSNIGHAQAAAGVAYIX 29955
QY 361	MMALRHGLRLPRLTHVDESPPHVMSAGAVOLLTETVWPGAGEGLRLRAGVSSFGVSGTN 420
Db 2996	MMALNLHGLRLPLSHIDESPPIHMSGCVRLTETVPMWQ-QNGRPRAAGVSAFGVSGTN 30545
QY 421	AHVILEEAP 429
	:

DB	3055	AHVIEQAP	3063
	RESULT 9		
XX	AA65266		
ID	AA65266	standard; Protein, 4881 AA.	
XX	AA65266;		
AC	04-DEC-2001	(first entry)	
DT			
XX	Streptomyces avermiltis protein SEQ ID NO: 6.		
DE			
XX	Avermectin aglycone synthase; AAs: avermectin derivative;		
KW	drug production; veterinary drug; pesticide.		
XX			
XX	Streptomyces avermiltis.		
OS			
XX	MO200162939-A1.		
PN			
XX	30-AUG-2001.		
PD			
XX	23-FEB-2001; 2001WO-JP01381.		
PF			
XX	24-FEB-2000; 2000JP-0047405.		
XX			
XX	(KYOW) KYOWA HAKKO KOGYO KK.		
PA	(KITA) KITASATO INST.		
XX			
PI	Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;		
XX	WPI; 2001-582053/65.		
DR	N-PSDB; AAH79278.		
XX			
PT	New modified avermectin aglycone synthase derived from Streptomyces		
PT	avermectilis used in production of 22,23-dihydroavermectin B1a used in		
PT	drugs and pesticides -		
XX			
XX	Claim 4; Page 201-217; 257pp; Japanese.		
XX			
XX	The present invention relates to the production of modified derivatives		
CC	of avermectin aglycone synthase (AAs) derived from Streptomyces		
CC	avermiltis. The activity of an acyl carrier protein (ACP),		
CC	beta-petoacyl carrier protein synthase (KS), acyltransferase (AT),		
CC	beta-petoacyl carrier protein reductase (KR), dehydratase (DH), enoyl		
CC	reductase (ER) and/or thioesterase (TE) domain may be reduced or		
CC	suppressed. The process can be used in the production of drugs, veterinary		
CC	drugs and pesticides. The present sequence is an S. avermiltis protein.		
XX			
XX	Sequence 4881 AA;		
XX			

Query Match	77.0%;	Score 1744.5;	DB 22;	Length 4881;
Match Local Similarity	76.7%;	Pred. No. 5.3e-146;		
Batches 329;	Conservative 40;	Mismatches 59;	Indels 1;	Gaps 1;
Qy	1	EPPIAVMACPFPGCVTSADPFMDLISSEODALGFPETDRGMDLDTLVPDPDPHGTCYT	60	
Db	2636	EPPIAVMACPFPGCVSAEDPFMLINSGRDAVEFPVDRQMDVEAFDPPEPGAGTSYT	26396	
Qy	61	RNGGFLYDAGHFDAEPFGISPRELAMDPOORLLLETFAMEETIEAGINPHTLHGTPTGVF	120	
Db	2696	RCGGFLGAAEFDFGFGISPRELAMDPOORLMLLEVSWEALIEAGIDPALTHTSTGVF	27555	
Qy	121	TGTNGDDVYALRVNHAAGSTDGFALTAGAGVISGRISYTFEGEPAYSVDPACSSLYAL	180	
Db	2756	AGVSGDDYAEILRRGTODHESGYALTGVSNVSYSRLSYTFEGEPAYTVDTACSSLYAL	28151	
Qy	181	HLACQALRAGECSMYALGCVTVMSPGAFAVEFSRQRIAGLADGHCKAESAADGTGMGEGV	240	
Db	2816	HLACQALRSGECSLALAGCVTVNSTPGAFAVEFSRQRIAGLADGHCKAESAADGTGMGAGV	28757	
Qy	241	GMLIVERYLSDAHRNGHGVYLAVRGSAVYNOGASNGLTAPNPGSGOQRIYRQLANAGLSAG	300	

Db 2876 GVLLEVERLSAERRGRHVLAVRGSAVNOGASNGLTAPNGSOORVIRIOLACAGLSVA 2935
 Qy 301 DVADEAAGCTGTTGDPDIEAQAALLATYAGDPAEGPMLGSKSNVGHOTQAAAGVAVIX 360
 Db 2936 DADVVEGHGCTTTGDDPIEAQALLATYGGGSGSRPWLGSVKSNIGHAQAAGVAVIX 2995
 Qy 361 MMALRLHGLLPTLHVDSPSPHVMASAGAVQLTETVPMWPGEGRLRAGVSGVSGTN 420
 Db 2996 MMALNHELLPTSLHIDESPPIHMSGSGVRLTLETPWQ-QNGRPRAAGVSAFGVSGTN 3054
 Qy 421 AHVLEBAP 429
 Db 3055 AHVIEQAP 3063
 RESULT 10
 AAY39299
 ID AAY39299 standard; Protein; 3170 AA.
 AC AAY39299;
 XX
 DT 01-DEC-1999 (first entry)
 XX
 DE SpnC a polyketide synthase.
 XX
 Km Spinosyn biosynthetic enzyme; open reading frame; ORF; insecticidal
 Km microtides; arachnid; nematode; insect; polyketide; polyketide synthase;
 Km PKS; extender module; initiator module; acyl transferase domain; AT;
 Km acyl carrier protein; ACP; beta-ketosynthase domain; KS; KR;
 Km dehydratase domain; DH; enoyl reductase domain; ER; beta-ketoreductase;
 Km insecticide.
 XX
 OS Saccharopolyspora spinosa.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..423
 FT /label= KS3
 FT /note= "Beta-ketosynthase domain: part of extender
 FT module 3"
 FT Domain 531..850
 FT /label= AT3
 FT /note= "Acyl transferase domain: part of extender
 FT module 3"
 FT Domain 1159..1337
 FT /label= KR3
 FT /note= "Beta-ketoreductase domain: part of extender
 FT module 3"
 FT Domain 1425..1506
 FT /label= ACP3
 FT /note= "Acyl carrier protein domain: part of extender
 FT module 3"
 FT Domain 1529..1952
 FT /label= KS4
 FT /note= "Beta-ketosynthase domain: part of extender
 FT module 4"
 FT Domain 2066..2396
 FT /label= AT4
 FT /note= "Acyl transferase domain: part of extender
 FT module 4"
 FT Domain 2700..2880
 FT /label= KR4
 FT /note= "Beta-ketoreductase domain: part of extender
 FT module 4"
 FT Domain 2972..3053
 FT /label= ACP4
 FT /note= "Acyl carrier protein domain: part of extender
 FT module 4"
 PF MO9946387-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 16-FEB-1999; 99WO-US03212.

XX 09-MAR-1998; 98US-0036387.
 PR (DOWC) DOW AGROSCIENCES LLC.
 PA
 XX Bultz RH, Broughton MC, Crawford KP, Madduri K, Merlo DJ;
 PI Treadway PJ, Turner JR, Waldron C;
 XX
 DR WPI; 1999-551414/46.
 DR N-PSDB; AA221501.
 XX
 PT New spinosyn biosynthetic genes from Saccharopolyspora spinosa, useful
 PT for production of insecticidal spinosyn compounds
 XX
 PS Claim 1; Page 103-113; 190pp; English.
 CC This is the amino acid sequence of the product of the spnC gene. The
 CC protein is involved in spinosyn biosynthesis. The spnC gene is one of 23
 CC genes and open reading frames contained in an 80kb DNA sequence
 CC AA221501. Spinosyns are insecticidal microtides which are useful for the
 CC control of arachnids, nematodes and insects. Biosynthesis of spinosyns
 CC occurs via stepwise condensation and modification of carboxylic acid
 CC precursors generating a linear polyketide which is modified further. The
 CC DNA sequence contains a central region of approximately 55kb which has
 CC homology to the DNA encoding the polyketide synthases (PKS) of known
 CC macrocyclic producers. The spinosyn PKS DNA region consists of 5 ORFs with
 CC stop codons at the end of acyl carrier protein (ACP) domains. Together
 CC the PKS polypeptides (AAY39297-39301) of which this sequence is one,
 CC form a complex consisting of an initiator module, spnA, and several
 CC extender modules spnB-spnH. Each extender module adds a specific acetyl
 CC Co-A precursor to a growing polyketide chain, and modifies the beta-keto
 CC group in a specific manner. A module in a PKS polypeptide consists of
 CC several domains with specific functions. The initiator module has an
 CC acyl transferase (AT) domain, and an acyl carrier protein (ACP) domain.
 CC The extender modules have the same domains plus a beta-ketosynthase (KS)
 CC domain, and optionally a beta-ketoreductase domain, a dehydratase (DH)
 CC domain, and an enoyl reductase (ER) domain. The last extender module
 CC terminates with a thiolester domain. The products of the genes present
 CC in the upstream region the PKS genes have been assigned names spnA-spnH
 CC AAY39302-39315 and are responsible for different modifications in
 CC spinosyn biosynthesis. There are also two ORFs ORF15 and ORF16 present
 CC immediately upstream of spnH, producing polypeptides AAY39316-39317,
 CC and two ORFs ORF1 and ORF2 present downstream of the PKS region
 CC producing polypeptides AAY39318-39319. The genes are useful to improve
 CC yields of spinosyns, and for creating new spinosyns e.g. by mutagenesis,
 CC or interruption of steps in spinosyn biosynthesis. The modified
 CC spinosyns may be a new insect control agent or serve as substrates for
 CC further chemical modification and the creation of new semi-synthetic
 CC spinosyns. The genes are also useful to isolate similar sequences from S.
 CC spinosa or other species by hybridization.
 XX
 SQ Sequence 3170 AA;
 Query Match 75.2%; Score 1704; DB 20; Length 3170;
 Best Local Similarity 75.3%; Pred. No. 1,2e-142;
 Matches 323; Conservative 38; Mismatches 66; Indels 2; Gaps 1;
 Qy 1 EPIAIVGMACRFGVTSADDFWDLISSEDAIGGPTDRGMWLDLYDPDPHPTGCTY 60
 Db 34 EPIAIVAMGCRYPGVQDEEGMKLVASGDAIGEPFARGMILDELYDDPDPQPGCTY 93
 Qy 61 RNSGPIYDAGHPAEFGISPREALAMPDPOORLLTFAMETIHAGINPTLTGTPGVV 120
 Db 94 RHGGFLHDAGEPFAGFDISPREALAMPDPOORLLTSMETTSAGNDPPLSGRTGVF 153
 Qy 121 TGTNGDYALRVHNAQSTDGAFALTGTAGSVISGRISYTFGFGPAVSVDTACSSSLVAL 180
 Db 154 AGLMVGYDTGAHRAAGEGVGLGTGNAGSVASGRAVAVAFGFGPAVTVDTCASSSLVAL 213
 Qy 181 HLAQOLRAGECGMALAGVTTWSSPGAFVPSRQGLAADGCKAFSAADGTGMEGV 240
 Db 214 HLAQOSLRQECDLALAGVTWSTPERFVPSRQGLAPDGRCKSFSAADGTGMEGA 273

XX This invention describes a novel method nucleic acid (I) and its encoded
 CC polypeptide (II) containing at least one region that encodes an enzymatic
 CC activity involved in biosynthesis of spinosyn. (I) are used (i) to
 CC identify, inactivate or modulate genes involved in the biosynthesis of
 CC (II), (ii) to generate a library of polypeptide synthetases, (iii) for
 CC adding fosfomycin or trimethylamine to a spinosyn or polypeptide
 CC synthetase, and (iv) for recombinant production of the corresponding
 CC enzymes, which are used for production of (II), their precursors or
 CC derivatives, including production of transgenic plants that express (II)
 CC and thus have increased resistance to insects. (i) are also useful as
 CC markers for sequencing of the *Saccharopolyspora spinosa* genome. (ii) are
 CC microtubules with insecticidal, but not antibacterial, activity, and can
 CC also be used to raise specific antibodies, useful for identifying
 CC expression clones in a gene bank. Cells transformed with (I) may produce
 CC (II) at significantly increased levels or produce new derivatives of
 CC (II). This sequence represents an *S. spinosa* polypeptide synthetase.

XX Sequence 4924 AA;

Query Match 73.8%; Score 1672; DB 22; Length 4924;

Best Local Similarity 75.2%; Pred. No. 1.6e-139; Matches 322; Conservative 34; Mismatches 70; Indels 2; Gaps 2;

QY 3 IAIYVMACTPPGVTSADDFMDLISSEDDAIGFPTDRGMDLTLYPDHPDPCYTRN 62
 DB 1783 VAIIYMGCRPPGVSPPEBLRLVAGVDVAGPPDRGMDLALYDPDRIGTSYVCE 1842
 QY 63 GGFPLVDAGHDAEPFGISPREALAMPQORLLLETAMETIEHAGINHTLHGPTGPTG 122
 DB 1843 GGFPLVDAGHDAEPFGISPREALAMPQORLLLETAMETIEHAGINHTLHGPTGPTG 1902
 QY 123 TNGDVALR-VHNAGSTDFALGTAGSVISGRISYTPGEPGPAVSDTACSSSLVALH 181
 DB 1903 LNHVDYARITTAPEGFEGHGLTGNGSVLSGRVAISFGEPGPAVSDTACSSSLVALH 1962
 QY 182 LACQALRAGSCSMALAGSVTVSSPGAFFVFSRQGLAADGHCFAFAAADGTGMBGVG 241
 DB 1963 LACQALRAGSCCEPALAGSVTVSSPTTFVFEFSRQGLAADGHCFAFAAADGTGMBGVG 2022
 QY 242 MLVVERSDAHNRHRLAVVAVRGAVNODGASNLTPNGPSQORVIRALANGLSAGD 301
 DB 2023 LVLLERISDARNGHEVLAVVAVRGAVNODGASNLTPNGPSQORVIRALANGLSAGD 2082
 QY 302 VDAVEAHGTGTTGADPIEAQALATATYGDRAEGEPLMGYSKSVNGHTOAAAGAVITKM 361
 DB 2083 VDAVEAHGTGTTGADPIEAQALATATYGDRAEGEPLMGYSKSVNGHTOAAAGAVITKM 2142
 QY 362 VVALRHGLPRTLHVDEPSPHVDMSAGAVQLTETVPWPGEGRLRAGVSSFGVSGTNA 421
 DB 2143 VVALRHGLPRTLHVDEPSPHVDMSAGAVQLTETVPWPGEGRLRAGVSSFGVSGTNA 2201
 QY 422 HVLLERAP 429
 DB 2202 HVLLERAP 2209

RESULT 13

AA139300 standard; Protein; 4928 AA.

AA139300;

01-DEC-1999 (first entry)

Spnd a polypeptide synthase.

Spinosyn biosynthetic enzyme; open reading frame; ORF; insecticidal
 microtubules; arachnid; nematode; insect; polypeptide; polypeptide synthase;
 PKG; extender module; initiator module; acyl transferase domain; AT;
 acyl carrier protein; ACP; beta-ketosynthase domain; KS; KS;
 dehydratase domain; DH; enoyl reductase domain; ER; beta-ketoreductase;
 insecticide.

XX Saccharopolyspora spinosa.
 OS
 XX Location/Qualifiers
 FH Key
 FT Domain
 FT 1..424
 FT /label= KS5
 FT /note= "beta-ketosynthase domain: part of extender
 module 5"
 FT Domain
 FT 539..866
 FT /label= AT5
 FT /note= "acyl transferase domain: part of extender
 module 5"
 FT Domain
 FT 893..1078
 FT /label= DH5
 FT /note= "dehydratase domain: part of extender
 module 5"
 FT Domain
 FT 1384..1565
 FT /label= KS5
 FT /note= "beta-ketoreductase domain: part of extender
 module 5"
 FT Domain
 FT 1645..1726
 FT /label= ACP5
 FT /note= "acyl carrier protein domain: part of extender
 module 5"
 FT Domain
 FT 1748..2172
 FT /label= KS6
 FT /note= "beta-ketosynthase domain: part of extender
 module 6"
 FT Domain
 FT 2283..2613
 FT /label= AT6
 FT /note= "acyl transferase domain: part of extender
 module 6"
 FT Domain
 FT 2916..3095
 FT /label= KR6
 FT /note= "beta-ketoreductase domain: part of extender
 module 6"
 FT Domain
 FT 3188..3269
 FT /label= ACP6
 FT /note= "acyl carrier protein domain: part of extender
 module 6"
 FT Domain
 FT 3291..3713
 FT /label= KS7
 FT /note= "beta-ketosynthase domain: part of extender
 module 7"
 FT Domain
 FT 3825..4153
 FT /label= AT7
 FT /note= "acyl transferase domain: part of extender
 module 7"
 FT Domain
 FT 4344..4638
 FT /label= KR7
 FT /note= "beta-ketoreductase domain: part of extender
 module 7"
 FT Domain
 FT 4725..4806
 FT /label= ACP7
 FT /note= "acyl carrier protein domain: part of extender
 module 7"
 XX WO946387-A1.
 XX 16-SEP-1999.
 XX 16-FEB-1999; 99WO-US03212.
 XX 09-MAR-1998; 98US-0036987.
 XX (DOWC) DOW AGROSCIENCES LLC.
 XX Balitz RH, Broughton MC, Crawford KP, Madduri K, Merlo DJ;
 PI Treadway PJ, Turner JR, Waldron C;
 DR WPI, 1999-551414/46.
 DR N-PSDB; AA221501.
 XX

PT New spinosyn biosynthetic genes from *Saccharopolyspora spinosa*, useful
PT for production of insecticidal spinosyn compounds -
XX
PS Claim 1; Page 113-129; 190pp; English.

QY	422	HVILEEAP	429
		: : :	
Db	2206	HLLEQPP	2211

RESULT 14
AAE10129
ID AAE10129 standard; Protein; 11096 AA

DT 29-NOV-2001 (first entry)

DE Streptomyces noursei nystatin gene, NY8C.

KM Polyketide synthase; PKS; macrolide; PKS gene cluster.
KM antifungal; antibiotic; PKS type I.

OS *Streptomyces noursei*.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

```
FT /label= KS3_domain
```

FT	Domain	546..858
----	--------	----------

/note= "Acyltransferase (AT) domain"

```
FT /label= DH3_domain
DE /date= "Debut2003 (m)" domain"
```

FT	Domain	1381..1628
----	--------	------------

/ note= "Ketoreductase (KR) domain"

```
FT      /label= ACP3_domain
DE      /label= "Acyl domain" (ACP) domain
```

FT	Domain	1757..2180
----	--------	------------

/note= "Ketosynthase (KS) domain"

```
FT      /label= AT4 domain
cm      /note= "NewI+transformation (AT) domain"
```

Domain	2617..2818
FT	

/note= "Dehydratase (DH) domain"

```
FT /label= KR4_domain
TE /label= "Krebs cycle" (kb) domain
```

FT	Domain	3407..3480
----	--------	------------

/note= "Acyl carrier protein (ACP)

```
FT /label= KS5_domain
TE /name= Myotomothars (wc) domain
```

FT	Domain	4032..4346
TE		/12801 - 12801

```

FT      /note= "Acyltransferase (AT) domain
EM      4350 4551

```

```

FT      /label= DH5_domain .
EE      /note= "Doherty2000 (DH) domain"

```

FT	Domain	4953..5239
FT		/12607 - EDC 30000000

FT	/note= "Enoylreductase (ER) domain"
EM	E349 E405
Domain	

```

FT      /label= KR5_domain
EM      /noto= "KatoCrustace (KR) domain"

```

Domain	5528..5601
FT	/1-hol - ACNE domain
EM	.

FT	Domain	/note="AcyI carrier protein (ACP) c
5633	5045	

```

PT      /label = KS6_domain
EM      /note = "Ketoisomerase (KS) domain"

```

FT Domain 6166..6478
 /label= AT6_domain
 /note= "Acyltransferase (AT) domain"
 FT Domain 6492..6704
 /label= DH6_domain
 /note= "Dehydratase (DH) domain"
 FT Domain 7038..7281
 /label= KR6_domain
 /note= "Ketoreductase (KR) domain"
 FT Domain 7315..7388
 /label= ACP6_domain
 /note= "Acyl carrier protein (ACP) domain"
 FT Domain 7408..7831
 /label= KS7_domain
 /note= "Ketosynthase (KS) domain"
 FT Domain 7939..8253
 /label= AT7_domain
 /note= "Acyltransferase (AT) domain"
 FT Domain 8267..8470
 /label= DH7_domain
 /note= "Dehydratase (DH) domain"
 FT Domain 8812..9086
 /label= KR7_domain
 /note= "Ketoreductase (KR) domain"
 FT Domain 9120..9193
 /label= ACP7_domain
 /note= "Acyl carrier protein (ACP) domain"
 FT Domain 9214..9637
 /label= KS8_domain
 /note= "Ketosynthase (KS) domain"
 FT Domain 9758..10072
 /label= AT8_domain
 /note= "Acyltransferase (AT) domain"
 FT Domain 10086..10289
 /label= DH8_domain
 /note= "Dehydratase (DH) domain"
 FT Domain 10657..10904
 /label= KR8_domain
 /note= "Ketoreductase (KR) domain"
 FT Domain 10939..11012
 /label= ACP8_domain
 /note= "Acyl carrier protein (ACP) domain"
 XX MO200159126-A2.
 XX 16-AUG-2001.
 XX 08-FEB-2001; 2001WO-GB00509.
 XX 08-FEB-2000; 2000GB-0002840.
 PR 10-APR-2000; 2000GB-0008786.
 PR 14-APR-2000; 2000GB-0009387.
 XX (UYNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
 PA (SNTF) SINTEF STIFTELSEN IND TEK FORSK.
 PA (ALPH-) ALPHARMA AS.
 PA (SINV-) SINVENT AS.
 PA (DZIE/) DZIELEWSKA H.
 PA (ZOTC/) ZOTCHEV S B.
 PA (SEKU/) SEKUROVA O N.
 PA (FJAE/) FJAEVRIK E.
 PA (BRAU/) BRAUTASET T.
 PA (STRO/) STROM A R.
 XX Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR,
 PI Valla S, Ellingsen TE, Sletta H, Gulliksen O,
 XX WPI; 2001-557614/62.
 DR N-PSDB; AAD17184.
 XX New nystatin polyketide synthase polynucleotides and polypeptides,
 PT useful as antibiotics and antifungals -
 XX

PS Claim 15; Page 170-176; 266pp; English.
 XX The present invention relates to the cloning and sequencing of the gene
 CC cluster encoding a modular type I polyketide synthase (PKS) enzyme
 CC involved in the biosynthesis of the macrocyclic antibiotic nystatin.
 CC The nystatin PKS is useful as antifungal antibiotic. The present
 CC sequence is a PKS type I encoding Streptomyces noursei nystatin
 CC gene, NysC.
 XX
 SQ Sequence 11096 AA;
 Query Match 73.1%; Score 1656; DB 22; Length 11096;
 Best Local Similarity 73.6%; Pred. No. 1,4e-137;
 Matches 320; Conservative 40; Mismatches 67; Indels 8; Gaps 3;
 QY 1 EPRLVGMCRFGGVTSDADFDLISSEDAAGCPTDRGMVLDLTPDPHPCGYT 60
 DB DPLVVVGMCRFGGVTSPEDLWRLVDDGTDAITTPTRNGMDLNLXDEDPHFCTSYT 3560
 QY 61 RNGGFLYDAGHPFAEFPGISPREALAMPDQORLLLETAMETTBAGINPHLLHGTPTGVF 120
 DB RSGGFLHNGEFPDPAFCMSPREALATDSQRLLLSSEWAIERAGIDPLTLGSGATGVF 3620
 QY 121 TGTNGDDYALRVHNAQSTDFALVTGTAGSVISGRISYTFGEPRVSVDTACSSSLVAL 180
 DB AGWMSDYGSIL--GGEKEFGFGQGSAGSVASGRVSYALGFEGPAVTVDTCSSSLVAL 3678
 QY 181 HLAQALRAGECEMALAGVTWSSPGAPVPEFRORGLAADGICAFSAADGTGEGV 240
 DB HMAQALRAGECSLAAAGVTWSTPTVPEFRORGLAPDGSKAPFAEALDVGSEGV 3738
 QY 241 GMLTVERLSDAHNGHGVLAIVVGSANODGASNGLTAPRPSQORVIRQALNAGLSAG 300
 DB GILVTEROSDAVNGEILAVIRGSANODGASNGLTAPRPSQORVIRQALNAGLSSTA 3798
 QY 301 DVDVAEAGGTGTLGDPTEAQAALLATYGGDRAGEGFLMGSVKSNVGHTOAAAGVAVIK 360
 DB DVDVAEAGGTGTLGDPTEAQAALLATYGRDRDENPLMGSLKSNIGHTOAAAGVAVIK 3858
 QY 361 MWALRHGILLPRLTHNDESPHVDNAGAVOLLTEVPVPGSGRARRAGVSFGVSGTN 420
 DB MWAMRHGVLPLTHVDAPSSHVDMSVGAVELLTEQTAMP-ETGRVRRAGVSSFGISGTN 3917
 QY 421 AHVILE-----EAPA 430
 DB 3918 AHVIEQPALVESPA 3932
 Db
 RESULT 15
 ID AAY92709
 ID AAY92709 standard; Protein; 3519 AA.
 XX AAY92709;
 AC AAY92709;
 XX 29-AUG-2000 (first entry)
 DT 29-AUG-2000 (first entry)
 XX S. antibioticus 8,8a-deoxyoleandolide synthase III encoded by oleAIII.
 DE oleandomycin; oleandolide; polyketide synthase; oleAI; oleAII; oleAIII;
 KW PKS; type I; 8,8a-deoxyoleandolide synthase; modular; ketosynthase;
 KM acyl-transferase; acyl carrier protein; inactivated; polyketide;
 KM macrolactone; antibiotic; molitide; erythromycin.
 XX
 OS Streptococcus antibioticus.
 XX
 FH Key Location/Qualifiers
 FT Region 33..1564
 FT /label= Extender_module_5
 FT Domain 33..466
 FT /label= KS5_domain
 FT /note= "Ketosynthase domain"
 FT 569..915
 FT /label= AT5_domain
 FT

FT /note= "acyl transferase domain"
 FT 1199..1383
 FT /label= KR5_domain
 FT /note= "ketoreductase domain"
 FT 1478..1564
 FT /label= ACP5_domain
 FT /note= "acyl carrier protein domain"
 FT 1686..3218
 FT /label= Extender_module_6
 FT 1686..2115
 FT /label= KS6_domain
 FT /note= "ketosynthase domain"
 FT 2220..2566
 FT /label= ACP6_domain
 FT /note= "acyl transferase domain"
 FT 2855..3039
 FT /label= KR6_domain
 FT /note= "ketoreductase domain"
 FT 3132..3218
 FT /label= ACP6_domain
 FT /note= "acyl carrier protein domain"
 FT 3290..3519
 FT /label= TE_domain
 FT /note= "thiol esterase domain"
 FT /note= "thiol esterase domain"
 PN WO200026349-A2.
 PD 11-MAY-2000.
 XX 22-OCT-1999; 99WO-US24478.
 XX 29-OCT-1998; 98US-0106100.
 PR 16-FEB-1999; 99US-0120254.
 XX
 PA (KOSA-) KOSAN BIOSCIENCES INC.
 XX
 PI Belach MC, Shah SK, McDaniel R, Tang L;
 XX
 DR WPI: 2000-365602/31.
 DR N-PSDB; AAA09469.
 XX
 PT Recombinant DNA compound encoding oleandolide polyketide synthase for
 PT synthesizing polyketides comprising a coding sequence for a domain of a
 PT loading module or any one of extender modules
 XX
 PS Disclosure; Page 29; 86pp; English.
 XX
 CC The oleandolide polyketide synthase (PKS), also known as
 CC 8,8a-deoxyoleandolide synthase, is encoded by three open reading frames
 CC (ORF), designated oleaI, oleaII and oleaIII. The PKS is a type I
 CC "modular" enzyme, where each ORF encodes 2 extender modules and
 CC the first ORF also encodes the loading module. Each module is composed
 CC of at least a ketosynthase (KS), acyl-transferase (AT) and an
 CC acyl carrier protein (ACP) domain. The oleandolide PKS loading module
 CC contains an inactivated KS, called KS-Q, where Q is the abbreviation for
 CC glutamine, present instead of the active site cysteine required for
 CC activity. The large multifunctional PKS enzymes catalyze the biosynthesis
 CC of polyketide macrolactones through multistep pathways involving
 CC decarboxylative condensations between acylthioesters followed by cycles
 CC of varying beta-carbon processing activities. The macroide product of
 CC the PKS, 8,8a-deoxyoleandolide, is further modified by epoxidation and
 CC glycosylation to yield oleandomycin, an antibacterial polyketide. The
 CC invention concerns an isolated recombinant DNA compound, comprising a
 CC coding sequence for a domain of loading module or any one of extender
 CC modules 1-4 or 1-6, including an oleandolide PKS operably linked to a
 CC promoter. Also discussed are recombinant oleandolide PKS in which the
 CC module 1 KS domain is inactivated by deletion or other mutation. In
 CC particular, the inactivation is mediated by a change in the KS domain
 CC that renders it incapable of binding substrate (the KS1-Q mutation),
 CC rendered by mutation in the codon for the active site cysteine. The
 CC oleandolide PKS is useful for synthesizing polyketides, which are useful
 CC as antibiotics and mollusks. Heterologous expression of oleandolide PKS
 CC in host cells such as Streptomyces coelicolor and S. lividans is also

CC made possible. Unmodified oleandolide compounds can be provided to
 CC cultures of Saccharopolyspora erythraea and converted to the
 CC corresponding derivatives of erythromycins A-D.
 XX
 SQ Sequence 3519 AA;
 QY
 1 EPIAIVGMACRPFPGVTSADDFWDLISSEDDAIGGFPTDRGMIDLTPDPDPHPCYCT 60
 1687 EALAIITAMSCRFGGIDSPEDLWFLAEGRDVAAGLPEDRGMDLALYHDDPENPGTTYV 1746
 61 RINGFLYDAGHPFAEPFGISPREALAMPQORLLLETAMETIHPAGINPHTLTGTPGVF 120
 1747 REGAFRIYDAQPTAGGFGISPREALAMPQORLLLETSMELFRAIDPPTVATGTATGIF 1806
 QY 121 TGTNGODVALRVNAGOSTDGFALTGTAGSVISGRISYTFGFBGPAVSVDTACSSSLVAL 180
 Db 1807 IGAHQGYGDPDKRAPEVSAGVILTGTASAVLSGRISYTFGLEGPATVDTACSSSLVAL 1866
 QY 181 HIAQALRAGEGSMALAGVTVMSFGATVFEBSRQGLAADGCKAFSAADGTGEGEV 240
 Db 1867 HIAVQALRREGCSIAIAGVAVMSTPDATFVFEBSRQGMARDGCKAFSAADGTGEGEV 1926
 QY 241 GMLVRLSDPAHNGHRLVAVVGSVAVNODGASNGLTAPNGSPQORVIRQALNAGLSAG 300
 Db 1927 SLTLRLSDPARLGRVLAIVGSAVNODGASNGLAAPNGSPQORVIRQALNAGLSAG 1986
 QY 301 DVDVAEAGTGTTAGDPTEAOLALATYGDPRAGEGPMILGSVKSNGHTQAAAGVAVIK 360
 Db 1987 DVDVVEAHGTGTLGDPTEAOLALATYGGORAGGRVPMVLSVKSNGHTQAAAGVAVIK 2046
 QY 361 MMALRHGILLPRLTHYDERSPHYDMSAGAVOLLTEVPVPGSGGRIRACVSGFVSGTN 420
 Db 2047 MYLALRGVVPKTLHVDERSPHYDMSAGAVELLTEERPEERLRRAGISAFVSGTN 2106
 QY 421 AHVILEAPA 430
 Db 2107 AHVILEAPA 2116

Search completed: June 17, 2003, 13:02:32
 Job time : 22.0646 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 17, 2003, 12:53:02 ; Search time 3.78986 Seconds
(without alignments)
4458.180 Million cell updates/sec

Title: US-09-914-286-3_COPY_1979_2060

Perfect score: 407
Sequence: 1 LAHQTSADQRAALBLVRDH.....RLPVSIVFDHPKAVLHL 82

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	407	100.0	3626	2 Q9F779	Q9F7C9 streptomyc
2	407	100.0	3972	2 Q9S0R8	Q9S0R8 streptomyc
3	243	59.7	5532	2 Q9S0R4	Q9S0R4 streptomyc
4	240.5	59.1	3524	2 Q93H86	Q93H86 streptomyc
5	240	59.0	6797	2 Q9X993	Q9X993 streptomyc
6	238	58.5	3576	16 Q9X53	Q9X53 streptomyc
7	233	57.2	3939	2 Q93HJ3	Q93HJ3 streptomyc
8	233	57.2	6239	2 Q9S0R7	Q9S0R7 streptomyc
9	229.5	56.4	3352	2 Q93H83	Q93H83 streptomyc
10	229	56.3	1839	2 Q30765	Q30765 streptomyc
11	229	56.3	3739	2 Q9ZG14	Q9ZG14 streptomyc
12	229	56.3	3970	2 Q93H18	Q93H18 streptomyc
13	229	56.3	6048	2 Q93H87	Q93H87 streptomyc
14	229	56.3	9477	2 Q9L4X3	Q9L4X3 streptomyc
15	226	55.5	6146	2 Q93HJ5	Q93HJ5 streptomyc
16	225.5	55.4	3654	2 Q30766	Q30766 streptomyc

17	225	55.3	1562	2 Q9ZG13	Q9ZG13 streptomyc
18	224	55.0	1762	2 Q52546	Q52546 amycolatops
19	224	55.0	1763	2 Q52790	Q52790 amycolatops
20	224	55.0	3816	2 Q9K1V3	Q9K1V3 streptomyc
21	223	54.8	2100	2 Q30481	Q30481 streptomyc
22	223	54.8	4150	2 Q9K1V4	Q9K1V4 streptomyc
23	223	54.8	9510	2 Q93XN9	Q93XN9 streptomyc
24	222	54.5	6396	2 Q9K1D7	Q9K1D7 streptomyc
25	221	54.3	5069	2 Q52545	Q52545 amycolatops
26	221	54.3	5069	2 Q52789	Q52789 amycolatops
27	221	54.3	6420	2 P95914	P95914 streptomyc
28	221	54.3	11096	2 Q9L4W3	Q9L4W3 streptomyc
29	220	54.1	4840	2 Q93HJ4	Q93HJ4 streptomyc
30	219	53.8	3729	2 Q33956	Q33956 streptomyc
31	219	53.8	4685	2 Q93HJ2	Q93HJ2 streptomyc
32	218	53.6	3613	2 Q93HJ1	Q93HJ1 streptomyc
33	217	53.3	4881	2 Q9S0R3	Q9S0R3 streptomyc
34	216.5	53.2	1762	2 Q30480	Q30480 streptomyc
35	216	53.1	3413	2 Q54593	Q54593 amycolatops
36	216	53.1	3591	2 Q9K1E1	Q9K1E1 streptomyc
37	216	53.1	4151	16 Q53490	Q53490 mycobacteri
38	215.5	52.9	7525	2 Q9K1E0	Q9K1E0 streptomyc
39	215	52.8	5435	2 Q9L4X2	Q9L4X2 streptomyc
40	215	52.8	8563	2 Q54297	Q54297 streptomyc
41	214.5	52.7	10223	2 Q54296	Q54296 streptomyc
42	212	52.1	1728	2 Q54591	Q54591 amycolatops
43	212	52.1	2152	2 Q9ALM5	Q9ALM5 baccharopol
44	212	52.1	6260	2 Q54299	Q54299 streptomyc
45	211	51.8	10917	2 Q93HJ6	Q93HJ6 streptomyc

ALIGNMENTS

RESULT 1

Q9F779 PRELIMINARY; PRT; 3626 AA.
ID Q9F779
AC Q9F779;
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Avermectin polyketide synthase (Fragment).
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC12711;
RA Hong Y.-S., Lee J.-J.;
RT "Targeted Gene Disruption of the avermectin O-methyltransferase gene
RT and polyketide synthase gene from Streptomyces avermitilis.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275943; AAC09812.1; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantane attach.
DR Pfam; PF00658; Acyl_transf; 3.
DR Pfam; PF00109; ketoacyl-synt; 2.
DR Pfam; PF02801; ketoacyl-synt C; 2.
DR Pfam; PF00550; pp-binding; 2.
DR PROSITE; PSS0075; ACP DOMAIN; 2.
DR PROSITE; PSS0066; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PSS0012; PHOSPHOPANTHEINE; UNKNOWN_1.
FT NON_TER
SQ
SEQUENCE 3626 AA; 380557 MW; 6272P5F088C1ABD0 CRC64;

Query Match 100.0%; Score 407; DB 2; Length 3626;
Best Local Similarity 100.0%; Pred. No. 1.6e-34;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 LAHQTSADQRAALBLVRDHVAALRHADPKAIPQSFALGFDLTAVFNNLIKAT 60
|||||

Db 1946 LAHOTSADQRAALLLEVRDHVAALRHADPKAIAIPDQSFALGFDLSITAVEFRNLIKAT 2005
 QY 61 GLRLPVSIVFDHPPTPAKLAAYHL 82
 Db 2006 GLRLPVSIVFDHPPTPAKLAAYHL 2027

RESULT 2
 Q9S0R8 PRELIMINARY; PRT; 3972 AA.
 AC Q9S0R8;
 DT 01-MAY-2000 (TREMBLREL. 13, Created)
 DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLREL. 20, Last annotation update)
 DE Type I polyketide synthase AVES 1.
 GN AYDAL.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 CX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99380548; PubMed=10449723;
 RA Ikeda H., Nonomiyama T., Usami M., Omura S.;
 RT "Organization of the biosynthetic gene cluster for the polyketide
 RT anthelmintic macroide avermectin in Streptomyces avermitilis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
 DR EMBL: AB032367; BAA84474.1;
 DR InterPro: IPR001227; Ac_transferase.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantne_attach.
 DR InterPro: IPR000834; Zn_carbDpept.
 DR Pfam: PF00698; Acyl_transf; 3.
 DR Pfam: PF00109; ketoacyl-synt; 2.
 DR Pfam: PF02801; ketoacyl-synt_C; 2.
 DR Pfam: PF00550; pp-binding; 3.
 DR PROSITE: PS50075; ACP_DOMAIN; 3.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_2.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 2.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 3972 AA; 416852 MW; 2A393695B032B1C3 CRC64;

Query Match 100.0%; Score 407; DB 2; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 1,7e-34;
 Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LAHOTSADQRAALLLEVRDHVAALRHADPKAIAIPDQSFALGFDLSITAVEFRNLIKAT 60
 Db 1979 LAHOTSADQRAALLLEVRDHVAALRHADPKAIAIPDQSFALGFDLSITAVEFRNLIKAT 2038
 QY 61 GLRLPVSIVFDHPPTPAKLAAYHL 82
 Db 2039 GLRLPVSIVFDHPPTPAKLAAYHL 2060

RESULT 3
 Q9S0R4 PRELIMINARY; PRT; 5532 AA.
 AC Q9S0R4;
 DT 01-MAY-2000 (TREMBLREL. 13, Created)
 DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLREL. 20, Last annotation update)
 DE Type I polyketide synthase AVES 3.
 GN AYDAL.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 CX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99380548; PubMed=10449723;
 RA Ikeda H., Nonomiyama T., Usami M., Omura S.;

RT "Organization of the biosynthetic gene cluster for the polyketide
 RT anthelmintic macroide avermectin in Streptomyces avermitilis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
 DR EMBL: AB032367; BAA84478.1;
 DR InterPro: IPR001227; Ac_transferase.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantne_attach.
 DR InterPro: IPR000834; Zn_carbDpept.
 DR Pfam: PF00698; Acyl_transf; 3.
 DR Pfam: PF00109; ketoacyl-synt; 3.
 DR Pfam: PF02801; ketoacyl-synt_C; 3.
 DR Pfam: PF00550; pp-binding; 3.
 DR PROSITE: PS50075; ACP_DOMAIN; 3.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 3.
 DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 3.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 5532 AA; 575193 MW; EC7B2069D44A057 CRC64;

Query Match 59.7%; Score 243; DB 2; Length 5532;
 Best Local Similarity 59.8%; Pred. No. 1,1e-16;
 Matches 48; Conservative 5; Mismatches 28; Indels 0; Gaps 0;

QY 1 LAHOTSADQRAALLLEVRDHVAALRHADPKAIAIPDQSFALGFDLSITAVEFRNLIKAT 60
 Db 1756 LAHOTSADQRAALLLEVRDHVAALRHADPKAIAIPDQSFALGFDLSITAVEFRNLIKAT 1815
 QY 61 GLRLPVSIVFDHPPTPAKLAAYHL 82
 Db 1816 GLRLPVSIVFDHPPTPAKLAAYHL 1837

RESULT 4
 Q93H86 PRELIMINARY; PRT; 3524 AA.
 ID Q93H86;
 AC Q93H86;
 DT 01-DEC-2001 (TREMBLREL. 19, Created)
 DT 01-DEC-2001 (TREMBLREL. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLREL. 21, Last annotation update)
 DE Modular polyketide synthase.
 GN PTEA2.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 CX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ichikawa J., Hanamoto A., Takahashi C.,
 RA Shinoe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: Deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL: AB070949; BAB69304.1;
 DR InterPro: IPR001227; Ac_transferase.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR001995; Mitoch_carrier.
 DR InterPro: IPR003880; Pantne_attach.
 DR Pfam: PF00698; Acyl_transf; 2.
 DR Pfam: PF00109; ketoacyl-synt; 2.
 DR Pfam: PF02801; ketoacyl-synt_C; 2.
 DR Pfam: PF00550; pp-binding; 2.
 DR PROSITE: PS50075; ACP_DOMAIN; 2.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_2.
 KM Phosphopantetheine.
 SQ SEQUENCE 3524 AA; 366619 MW; 64A65759C461EB64 CRC64;

Query Match 59.1%; Score 240.5; DB 2; Length 3524;
 Best Local Similarity 65.4%; Pred. No. 1,2e-16;

Qy 1 LAHOTSADORAALLEVDHVAALRHADPKAIAPDGSFALGFDSTLAVFERNLLIKAT 60
 ID Q93HJ3 PRELIMINARY; PRT; 3939 AA.
 AC Q93HJ3;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Modular polyketide synthase.
 GN OLM43.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: Deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL: AB070940; BAB69194.1; -
 DR InterPro: IPR002105; AATRNA_ligaseII.
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR002085; Adh zn family.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantane attach.
 DR InterPro: IPR002364; OOR zeta crystal.
 DR Pfam: PF00698; Acyl_transf. 2.
 DR Pfam: PF00107; adh_zinc. 1.
 DR Pfam: PF00109; ketoacyl-synt. 2.
 DR Pfam: PF02801; ketoacyl-synt_C. 2.
 DR Pfam: PF00550; pp-binding. 2.
 DR PROSITE: PS00339; AA TRNA LIGASE II_2; UNKNOWN_1.
 DR PROSITE: PS50075; ACP_DOMAIN. 2.
 DR PROSITE: PS00606; B_KETOACYL SYNTHASE; UNKNOWN_2.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 DR PROSITE: PS01162; OOR_ZETA_CRYSTAL; UNKNOWN_1.
 KM Phosphopantetheine.
 SQ SEQUENCE 3939 AA; 410482 MW; 534AE3B517256503 CRC64;
 Query Match 57.2%; Score 233; DB 2; Length 3939;
 Best Local Similarity 61.3%; Pred. No. 8.4e-16;
 Matches 46; Conservative 9; Mismatches 20; Indels 0; Gaps 0;
 Qy 5 TSADORAALLEVDHVAALRHADPKAIAPDGSFALGFDSTLAVFERNLLIKAT 64
 ID Q93HJ3 PRELIMINARY; PRT; 6239 AA.
 AC Q93HJ3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Type I polyketide synthase AVS 2.
 DB 2069 PVTVPFDPHPTSELA 2083

GN AVEA2.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99380548; PubMed=10449723;
 RA Ikeda H., Nonomiya T., Usami M., Ohta T., Omura S.;
 RT "Organization of the biosynthetic gene cluster for the polyketide
 RT anheimitic macrolide avermectin in Streptomyces avermitilis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
 DR EMBL: AB032367; BAA84475.1; -
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantane attach.
 DR InterPro: IPR000169; Shprot_acsite.
 DR InterPro: IPR000834; Zn carbDopept.
 DR Pfam: PF00698; Acyl_transf. 4.
 DR Pfam: PF00109; ketoacyl-synt. 4.
 DR Pfam: PF02801; ketoacyl-synt_C. 4.
 DR Pfam: PF00550; pp-binding. 4.
 DR PROSITE: PS50075; ACP_DOMAIN. 4.
 DR PROSITE: PS00606; B_KETOACYL SYNTHASE; 4.
 DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_4.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 3.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 6239 AA; 666283 MW; A706280940B502FA CRC64;
 Query Match 57.2%; Score 233; DB 2; Length 6239;
 Best Local Similarity 61.0%; Pred. No. 1.4e-15;
 Matches 50; Conservative 8; Mismatches 24; Indels 0; Gaps 0;
 Qy 1 LAHOTSADORAALLEVDHVAALRHADPKAIAPDGSFALGFDSTLAVFERNLLIKAT 60
 ID Q93HJ3 PRELIMINARY; PRT; 3352 AA.
 AC Q93HJ3;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Modular polyketide synthase.
 GN PREA5.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: Deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL: AB070949; BAB69307.1; -
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR002362; Antenna_beta.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantane attach.
 DR InterPro: IPR001031; Thioesterase.

DR pfam; PF00698; Acyl_transf; 2;
 DR pfam; PF00106; adh_short; 1;
 DR pfam; PF00109; ketoacyl-synt; 2;
 DR pfam; PF02801; ketoacyl-synt_C; 2;
 DR pfam; PF00550; pp-binding; 2;
 DR pfam; PF00975; thioesterase; 1;
 DR PROSITE; PS50075; ACP_DOMAIN; 2;
 DR PROSITE; PS00669; ANTENNA_COMP_BETA; UNKNOWN_1;
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2;
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_2;
 KW Phosphopantetheine.
 SQ SEQUENCE 3352 AA; 353395 MW; 66A5865BFD2027B7 CRC64;

Query Match 56.4%; Score 229.5; DB 2; Length 3352;
 Best Local Similarity 58.5%; Pred. No. 1.7e-15;
 Matches 48; Conservative 10; Mismatches 21; Indels 3; Gaps 1;

QY 1 LAHQTSADQRAALLLEVRDHAVALRHADPKAIPDQSFALGFDLSITAVEFNRLIKAT 60
 DB 1426 LSHQ---DRARALVDLVRTTHAAVLGHRGAGAVEADRFDRIDGFDLSITAVEFNRLIKAT 1482
 61 GLRLPVSIVPDHPTPAKLAVAL 82
 DB 1483 GLSLPATLVFDHPADLAAL 1504

RESULT 10

030765 PRELIMINARY; PRT; 1839 AA.
 AC 030765;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Polyketide synthase module 3.
 GN NIDA2.
 OS Streptomyces caelestis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=36816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL-2821;
 RX MEDLINE=98053867; PubMed=9393718;
 RA Kakavae S.J., Katz L., Steasel D.;
 RT "Identification and characterization of the niddamycin polyketide
 RT synthase genes from Streptomyces caelestis.";
 RL J. Bacteriol. 179:7515-7522(1997).
 DR EMBL; AF016585; AAC46025.1;
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Pantane attach.
 DR pfam; PF00698; Acyl_transf; 1.
 DR pfam; PF00109; ketoacyl-synt; 1.
 DR pfam; PF02801; ketoacyl-synt_C; 1.
 DR pfam; PF00550; pp-binding; 1.
 DR PROSITE; PS50075; ACP_DOMAIN; 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 KW Phosphopantetheine; transferase.
 SQ SEQUENCE 1839 AA; 194158 MW; 46683C00526090A2 CRC64;

Query Match 56.3%; Score 229; DB 2; Length 1839;
 Best Local Similarity 60.8%; Pred. No. 9.3e-16;
 Matches 48; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 1 LAHQTSADQRAALLLEVRDHAVALRHADPKAIPDQSFALGFDLSITAVEFNRLIKAT 60
 DB 1662 LARTVABRLPLLELVRTTHAAVLGHRGAGAVEADRFDRIDGFDLSITAVEFNRLIKAT 1721
 61 GLRLPVSIVPDHPTPAKLA 79
 DB 1722 GLRLPATLVFDHPADLAAL 1740

RESULT 11

09ZG14 PRELIMINARY; PRT; 3739 AA.
 AC 09ZG14;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Type I polyketide synthase pikari.
 GN PIKARI.
 OS Streptomyces venezuelae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=45471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC15439;
 RX MEDLINE=98445333; PubMed=9770448;
 RA Xue Y., Zhao L., Liu H.W., Sherman D.H.;
 RT "A gene cluster for macroide antibiotic biosynthesis in streptomyces
 RT venezuelae: architecture of metabolic diversity";
 RT Proc. Natl. Acad. Sci. U.S.A. 95:12111-12116(1998).
 DR EMBL; AF079138; AAC69330.1;
 DR HSSP; Q02054; 2AF8.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh zn family.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Pantane attach.
 DR InterPro; IPR002364; OOR_zeta_crystal.
 DR InterPro; IPR00169; SHPOC_acsite.
 DR pfam; PF00698; Acyl_transf; 2.
 DR pfam; PF00107; adh_zinc; 1.
 DR pfam; PF00109; ketoacyl-synt; 2.
 DR pfam; PF02801; ketoacyl-synt_C; 2.
 DR pfam; PF00550; pp-binding; 2.
 DR PROSITE; PS50075; ACP_DOMAIN; 2.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
 DR PROSITE; PS01162; OOR_ZETA_CRYSTAL; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
 KW Phosphopantetheine; transferase.
 SQ SEQUENCE 3739 AA; 387170 MW; 3D3910824DA5B080 CRC64;

Query Match 56.3%; Score 229; DB 2; Length 3739;
 Best Local Similarity 59.8%; Pred. No. 2.1e-15;
 Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;

QY 1 LAHQTSADQRAALLLEVRDHAVALRHADPKAIPDQSFALGFDLSITAVEFNRLIKAT 60
 DB 3560 LAAMTPDDRVAHRLDLVRTTHAAVLGHRGAGAVEADRFDRIDGFDLSITAVEFNRLIKAT 3619
 61 GLRLPVSIVPDHPTPAKLAVAL 82
 DB 3620 GLRLPATLVFDHPADLAAL 3641

RESULT 12

093HI8 PRELIMINARY; PRT; 3970 AA.
 AC 093HI8;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Modular polyketide synthase.
 GN OLM44.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Iehikawa J., Hanamoto A., Takahashi C.,

RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: Deducing the ability of producing secondary
 RT metabolites.",
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL; AB070940; BAB69199.1; -.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; adh zn family.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantoate_attach.
 DR Pfam; PF00698; Acyl_transf. 4.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt. 2.
 DR Pfam; PF02801; ketoacyl-synt_C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PS50075; ACP_DOMAIN; 2.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 DR PROSITE; PS01162; OOR_ZETA_CRYSTAL; UNKNOWN_1.
 KW Phosphopantetheine.
 SQ SEQUENCE 3870 AA; 412657 MW; A93B583FAFA68C47 CRC64;
 Query Match 56.3%; Score 229; DB 2; Length 3970;
 Best Local Similarity 65.8%; Pred. No. 2.3e-15;
 Matches 50; Conservative 6; Mismatches 20; Indels 0; Gaps 0;
 QY 7 ADGPAALLETVDHVAALVLRHADPKAIAPOQSFALGFDSTLTAVEFNNLIKATGLRLPV 66
 DB 3811 AEOQRRLDLVRSNAATVLGHASAGAIABEQSFSELDITLTAVEFNNLIKATGLRLPA 3870
 QY 67 SLVFDHPPTPAKLAVHL 82
 DB 3871 TLVEHPPTPALAGEL 3886
 RESULT 13
 Q93887 ID Q93887 PRELIMINARY; PRT; 6048 AA.
 AC Q93887;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Modular polyketide synthase.
 GN PPEAL.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 ON NCBI_TaxID=33903;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: Deducing the ability of producing secondary
 RT metabolites.",
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL; AB070949; BAB69303.1; -.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR004410; Fabd.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantoate_attach.
 DR Pfam; PF00698; Acyl_transf. 4.
 DR Pfam; PF00109; ketoacyl-synt. 4.
 DR Pfam; PF00550; pp-binding; 4.
 DR TIGRfams; TIGR00128; fabd; 4.
 DR PROSITE; PS50075; ACP_DOMAIN; 4.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_3.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_4.

KW Phosphopantetheine.
 SQ SEQUENCE 6048 AA; 629420 MW; F8673C2ED2694705 CRC64;
 Query Match 56.3%; Score 229; DB 2; Length 6048;
 Best Local Similarity 57.3%; Pred. No. 3.7e-15;
 Matches 47; Conservative 11; Mismatches 24; Indels 0; Gaps 0;
 QY 1 LAHQTSADQRAALLEVRDHVAALVLRHADPKAIAPOQSFALGFDSTLTAVEFNNLIKAT 60
 DB 5878 LEEMPEAREKTYVDLVRAVAAGLGHASDRTVRPFAPODGLFDSITLTAVEFNNLIKAT 5937
 QY 61 GLRLPVSLVFDHPPTPAKLAVHL 82
 DB 5938 GLRLPATLVFDHPPTPALAHV 5959
 RESULT 14
 Q9L4X3 ID Q9L4X3 PRELIMINARY; PRT; 9477 AA.
 AC Q9L4X3;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE NysI.
 GN NYSI.
 OS Streptomyces noursei.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 ON NCBI_TaxID=19771;
 RX MEDLINE=20334850; PubMed=10873841;
 RA Brautaset T., Sekurova O.N., Sletta H., Ellingsen T.E., Strom A.R.,
 RA Valla S., Zocchev S.B.,
 RT "Biosynthesis of the polyene antifungal antibiotic nystatin in
 RT Streptomyces noursei ATCC 11455: analysis of the gene cluster and
 RT deduction of the biosynthetic pathway.",
 RL Chem. Biol. 7:395-403(2000).
 DR EMBL; AF263812; AAF71766.1; -.
 DR HSP; P81989; 10E5.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR004410; Fabd.
 DR InterPro; IPR001899; Gram_pos_anchor.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR001005; Myb_DNA_binding.
 DR InterPro; IPR003880; Pantoate_attach.
 DR Pfam; PF00698; Acyl_transf. 6.
 DR Pfam; PF00106; adh_short; 3.
 DR Pfam; PF00109; ketoacyl-synt. 6.
 DR Pfam; PF02801; ketoacyl-synt_C; 6.
 DR Pfam; PF00550; pp-binding; 6.
 DR TIGRfams; TIGR00128; fabd; 6.
 DR PROSITE; PS50075; ACP_DOMAIN; 6.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 6.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 6.
 KW Phosphopantetheine; Transferase.
 SQ SEQUENCE 9477 AA; 988132 MW; 0D2BCA5D7B265483 CRC64;
 Query Match 56.3%; Score 229; DB 2; Length 9477;
 Best Local Similarity 59.8%; Pred. No. 6.3e-15;
 Matches 49; Conservative 5; Mismatches 28; Indels 0; Gaps 0;
 QY 1 LAHQTSADQRAALLEVRDHVAALVLRHADPKAIAPOQSFALGFDSTLTAVEFNNLIKAT 60
 DB 7801 LAEQPADHAIILALVTEGAATVLRHADHDLEPDLPRDGLFDSITLTAVEFNNLIKAT 7860
 QY 61 GLRLPVSLVFDHPPTPAKLAVHL 82
 DB 7861 GLRLPATLVFDHPPTPALAHL 7882

RESULT 15

```

ID 093HJ5      PRELIMINARY;      PRT; 6146 AA.
AC 093HJ5.
DT 01-DEC-2001 (TRENBLREL. 19, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Modular polyketide synthase.
GN OLMA1.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinoue M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shida T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: Deducing the ability of producing secondary
RT metabolites."
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RL EMBL; AB070940; BAB69192.1;
DR InterPro; IPR002106; ALCRNA_LigaseII.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR004410; FAD.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR000209; Peptidase 58.
DR InterPro; IPR003880; Pantone attach.
DR InterPro; IPR002364; QOR_zeta_crystal.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00698; Acyl_transfer; 4.
DR Pfam; PF00107; adh_zinc; 1.
DR Pfam; PF00109; ketoacyl-synt; 4.
DR Pfam; PF02801; ketoacyl-synt; 4.
DR Pfam; PF00550; pp-binding; 4.
DR Pfam; PF00108; thiolase; 1.
DR TIGRfam; TIGR00128; fad; 4.
DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR PROSITE; PS00075; ACP_DOMAIN; 4.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_3.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
DR PROSITE; PS01162; QOR_ZETA_CRYSTAL; UNKNOWN_1.
DR PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
KW phosphopantetheine.
SQ SEQUENCE 6146 AA; 637460 MW; 1D56F1108D458D00 CRC64;

```

Query Match 55.5%; Score 226; DB 2; Length 6146;

Best Local Similarity 56.1%; Pred. No. 8e-15;

Matches 46; Conservative 10; Mismatches 26; Indels 0; Gaps 0;

```

QY 1 LAHQTSADQRALELVDRHVAALVRHADPKAIAPOGFRALGFDSLTAVERFRLIKAT 60
   |||:::|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 5980 LAASGPAERHRLVLDVREHRTAGVGHGSADIDPDQAFTLGFDSLTAVERLRRLKAT 6039
   |||:::|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 61 GLRLVSLVDPDPTPAKLAHL 82
   |||:::|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 6040 SLTVPATLVDPDPTPAKLAHL 6061
   |||:::|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

Search completed: June 17, 2003, 13:07:55
 Job time : 4.78986 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 12:52:22 ; Search time 0.894236 Seconds
(without alignments)
3803.313 Million cell updates/sec

Title: US-09-914-286-3_COPY_1979_2060
Perfect score: 407
Sequence: 1 LAHQTSADGRALLETVRDH.....RLPVSIVFDHPTAKLAVHL 82

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	205	50.4	3172	1 ERY3_SACER	003133 saccharopol
2	205	50.4	3567	1 ERY2_SACER	003132 saccharopol
3	202	49.6	3491	1 ERY1_SACER	003131 saccharopol
4	186	45.7	3519	1 OL56_STRAT	007017 streptomyc
5	104.5	25.1	4427	1 PKSL_BACSU	005470 bacillus su
6	98	24.1	1538	1 PRSB_MYCTU	Q10978 mycobacteri
7	94.5	23.2	1876	1 PPSA_MYCTU	Q10977 mycobacteri
8	87	21.4	1278	1 DMBF_BACSU	P45745 bacillus su
9	82.5	20.3	1774	1 MSAS_PENPA	P22367 penicillium
10	77	18.9	2110	1 MCAS_MYCBO	002251 mycobacteri
11	75.5	18.6	1392	1 LYS2_YEAST	P07702 saccharomyc
12	71.5	17.6	2505	1 PAS_FAT	P12785 rectus nov
13	71	17.4	92	1 ACP_STRCO	Q97K61 streptomyc
14	70.5	17.3	5217	1 HTS1_COCCA	Q01886 cochlidiol
15	70	17.2	3587	1 SRF1_BACSU	P27206 bacillus su
16	69	17.0	6486	1 TYCC_BACBR	Q30409 b tyrocidin
17	68.5	16.8	838	1 PAS_MOUSE	P19096 mus musculu
18	68.5	16.8	3649	1 ACVS_NOCIA	P27743 nocardi
19	68	16.7	5255	1 BACA_BACLI	Q68006 b bacitraci
20	67.5	16.6	3770	1 ACVS_EMBE1	P27742 emericella
21	66.5	16.3	1274	1 SRF3_BACSU	Q08787 bacillus su
22	66	16.2	2035	1 HMP2_YEREN	P48633 yeastoma en
23	65	16.0	88	1 ACPX_STRCM	P41174 streptomyc
24	65	16.0	2560	1 PRS2_BACSU	P39846 bacillus su
25	65	16.0	3587	1 TYCB_BACBR	Q30408 b tyrocidin
26	64.5	15.8	122	1 ACPM_ARATH	P33665 arabidopsi
27	64.5	15.6	357	1 DP42_AGRIS	Q04171 arabidopsi
28	63.5	15.6	568	1 DNLI_METAC	Q08121 methanococ
29	63.5	15.6	385	1 CRV2_ARATH	Q96554 arabidopsi
30	63	15.5	365	1 YL90_MYCTU	Q10389 mycobacteri
31	62.5	15.4	332	1 GALR_HAEIN	P31766 haemophilus
32	62	15.2	295	1 ISPE_CACCR	Q04817 caulobacter
33	62	15.2	420	1 GATZ_ECOLI	P37191 escherichia

ALIGNMENTS

RESULT 1
ID ERY3_SACER STANDARD: PRT; 3172 AA.
AC 003133: 054097: 099270:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Erythronolide synthase, modules 5 and 6 (EC 2.3.1.94) (ORF 3) (6-deoxyerythronolide B synthase III) (DEBS 3).
GN ERYA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
OC Actinomycetales; Pseudonocardiales; Pseudonocardaceae;
OC Saccharopolyspora.
OX NCBI_TaxID=1836;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RX MEDLINE=91043075; PubMed=2234082;
RA Cortes J., Haydock S.F., Roberts G.A., Bevil D.J., Leadlay P.F.;
RT "An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea.";
RT Nature 348:176-178(1990).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RX MEDLINE=92152065; PubMed=2024119;
RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT "Modular organization of genes required for complex polyketide biosynthesis.";
RT Science 252:675-679(1991).
RL [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RX MEDLINE=92152230; PubMed=1740151;
RA Bevil D.J., Cortes J., Haydock S.F., Leadlay P.F.;
RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme.";
RT Eur. J. Biochem. 204:39-49(1992).
RL [4]
RP CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-deoxyerythronolide B.
CC -1- COFACTOR: NADP; CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETHEINES.
CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACTL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR), DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION

OF THE FULL-LENGTH CHAIN.
 -1- SIMILARITY: TO FATY ACID SYNTHASE (FAS).
 -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch>).

 EMBL; X56107; CA39583.1; -
 EMBL; X63677; AAA26495.1; -
 EMBL; X62569; CAA44449.1; -
 HSSP; P00101; ICCH.
 InterPro; IPR002198; ADH_short.
 InterPro; IPR001227; Ac_transferrase.
 InterPro; IPR000794; Ketoacyl-synt.
 InterPro; IPR003880; Pantne attach.
 InterPro; IPR001031; Thioesterase.
 Pfam; PF00106; adh_short; 1.
 Pfam; PF00109; ketoacyl-synt; 2.
 Pfam; PF00550; pp-binding; 2.
 Pfam; PF00698; Acyl_transf; 2.
 Pfam; PF00975; Thioesterase; 1.
 Pfam; PF02801; ketoacyl-synt C; 2.
 PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
 PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
 PROSITE; PS50075; ACP DOMAIN; 2.
 Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 Phosphopantetheine; Multifunctional enzyme.

 DOMAIN 1 1484
 1 1484
 1485 3172
 37 484
 554 878
 1116 1298
 1394 1464
 1488 1954
 2021 2335
 2255 2735
 2821 2891
 2926 3172
 199 199
 643 643
 1118 1164
 1427 1427
 1661 1661
 2112 2112
 2557 2605
 2854 2854
 231 231
 240 240
 289 289
 493 493
 517 517

 MODULE 5.
 BETA-KETOACYL SYNTHASE 1.
 ACYLTRANSFERASE (AT) 1.
 BETA-KETOACYL REDUCTASE 1.
 ACYL CARRIER (ACP) 1.
 BETA-KETOACYL SYNTHASE 2.
 ACYLTRANSFERASE (AT) 2.
 BETA-KETOACYL REDUCTASE 2.
 ACYL CARRIER (ACP) 2.
 THIOESTERASE.
 THIOESTER BOND.
 ACYL-ENZYME INTERMEDIATE.
 NADP.
 PHOSPHOPANTETHEINE (BY SIMILARITY).
 THIOESTER BOND.
 ACYL-ENZYME INTERMEDIATE.
 NADP.
 PHOSPHOPANTETHEINE (BY SIMILARITY).
 S -> SA (IN REF. 2).
 MISSING (IN REF. 2).
 A -> R (IN REF. 2).
 P -> R (IN REF. 2).
 PEPRSLDGTGTLATRASAMEHRA -> ASRGTRCATPVS
 RMPPALAPWEO (IN REF. 1).
 A -> R (IN REF. 2).
 M -> W (IN REF. 2).
 E -> D (IN REF. 2).
 R -> G (IN REF. 2).
 GPNSP -> ARTR (IN REF. 2).
 R -> G (IN REF. 2).
 R -> A (IN REF. 2).
 MISSING (IN REF. 2).
 AHK -> GIT (IN REF. 2).
 R -> ROLPVPFOROR (IN REF. 1).
 GVAAPH -> VSLSD (IN REF. 2).
 RTHLEPLA -> ARTKSPR (IN REF. 2).
 MISSING (IN REF. 1).
 L -> V (IN REF. 2).
 A -> R (IN REF. 2).

FT CONFLICT 1194 1194 MISSING (IN REF. 2).
 FT CONFLICT 1277 1278 AA -> RR (IN REF. 2).
 FT CONFLICT 1385 1390 LCGRE -> STARR (IN REF. 2).
 FT CONFLICT 1485 1485 MISSING (IN REF. 2).
 FT CONFLICT 1518 1518 G -> R (IN REF. 2).
 FT CONFLICT 1601 1601 V -> L (IN REF. 2).
 FT CONFLICT 1724 1725 LP -> PA (IN REF. 2).
 FT CONFLICT 1732 1732 O -> L (IN REF. 2).
 FT CONFLICT 1739 1743 GPAEG -> ARRA (IN REF. 2).
 FT CONFLICT 1762 1762 T -> S (IN REF. 2).
 FT CONFLICT 2252 2252 D -> DQAD (IN REF. 2).
 FT CONFLICT 2275 2277 QSP -> AVA (IN REF. 2).
 FT CONFLICT 2408 2408 G -> GR (IN REF. 2).
 FT CONFLICT 2420 2421 LA -> S (IN REF. 2).
 FT CONFLICT 2443 2444 NA -> TH (IN REF. 2).
 FT CONFLICT 2596 2596 A -> G (IN REF. 2).
 FT CONFLICT 2609 2609 P -> A (IN REF. 2).
 FT CONFLICT 2715 2722 RAEGRRA -> AVRAVRR (IN REF. 1).
 FT CONFLICT 2754 2754 D -> E (IN REF. 2).
 SQ SEQUENCE 3172 AA; 331474 MW; DBD5094E77DD5F CRC64;

Query Match 50.4%; Score 205; DB 1; Length 3172;
 Best Local Similarity 58.6%; Pred. No. 3e-14;
 Matches 41; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

Qy 13 LLELVDAVAVRHADPKAIADQSFALGPSLTAVERNLLIKATGLRLPVSLVFDH 72
 Db 2821 LLEFTHVAAILGHSPDPAVGODPTELGPSLTVGRNQLQCATGALATVFEH 2880
 Qy 73 PTPAKLAHL 82
 Db 2881 PTVRLADHI 2890

RESULT 2
 ERY2_SACER STANDARD; PRT; 3567 AA.
 ID ERY2_SACER
 AC Q03132; O54096;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Erythronolide synthase, modules 3 and 4 (EC 2.3.1.94) (ORF 2) (6-deoxyerythronolide B synthase II) (DEBS 2).
 GN ERYA.
 OS Saccharopolyspora erythraea (Streptomyces erythraeus).
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
 OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae;
 OC Saccharopolyspora.
 OX NCBI_TaxID=1836;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220065; PubMed=2024119;
 RA Donadio S., Staver M.O., McAlpine J.B., Swanson S.J., Katz L.;
 RT "Modular organization of genes required for complex polyketide biosynthesis.";
 RL Science 252:675-679 (1991).
 RP [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=NRRL 2338;
 RX MEDLINE=92155230; PubMed=1740151;
 RA Bevil D.J., Cortes J., Haydock S.F., Leadlay P.F.;
 RT "6-deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme.";
 RL Eur. J. Biochem. 204:39-49 (1992).
 RL -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-deoxyerythronolide B.
 CC -1- COFACTOR: NADP. CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.
 CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
 CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH

CC SYNTASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS
CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
CC RESPECTIVELY.
CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),
CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACTYL CARRIER
CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
CC OF THE FULL-LENGTH CHAIN.
CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
CC -1- SIMILARITY: CONTAINS 2 ACTYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63677; AAA26494.1; -;
DR EMBL; X62569; CAA44448.1; -;
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002085; Adh_zn family.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Pantane_attach.
DR Pfam; PF00109; ketoacyl-synt; 2.
DR Pfam; PF00550; pp-binding; 2.
DR Pfam; PF00698; Acyl_transf; 2.
DR Pfam; PF02801; ketoacyl-synt C; 2.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
DR PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
DR PROSITE; PS50075; ACP_DOMAIN; 2.
DR TransErase; Acyltransferase; Antibiotic biosynthesis; NADP. Repeat;
KW Phosphopantetheine; Multifunctional enzyme.
FT DOMAIN 1 1484
FT DOMAIN 1485 3567
FT DOMAIN 27 488
FT DOMAIN 559 884
FT DOMAIN 1130 1301
FT DOMAIN 1397 1467
FT DOMAIN 1485 1943
FT DOMAIN 2013 2336
FT DOMAIN 2383 3066
FT DOMAIN 3139 3322
FT DOMAIN 3415 3485
FT ACT_SITE 202 202
FT ACT_SITE 651 651
FT BINDING 1430 1430
FT ACT_SITE 1661 1661
FT ACT_SITE 2115 2115
FT NP_BIND 2961 2978
FT NP_BIND 3142 3157
FT BINDING 3438 3448
FT CONFLICT 438 438
FT CONFLICT 480 480
FT CONFLICT 1241 1241
FT CONFLICT 2664 2664
SQ SEQUENCE 3567 AA; 374413 MW; EE6284F4738A0C0 CUC64;
Query Match 50.4%; Score 205; DB 1; Length 3567;
Best Local Similarity 53.7%; Pred. No. 3,4e-14;
Matches 44; Conservative 11; Mismatches 27; Indels 0; Gaps 0;
QY 1 LAHOTSADQRAALTELVRDVAVALRHADPKAIPDPSFALGDSLTAVEFRLLTKAT 60
DB 3403 LAGRSBEDQVAGLAEVLRSHAAVSGSGADQLPRKAFKDLGDSLAALVELNRRLGTAT 3462
QY 61 GLRLPVLVFDHPTAKLAVHL 82

DB 3463 GVALPSTLVFDHPTAKLAVAEHL 3484
RESULT 3
ID ERYL_SACER STANDARD; PRT; 3491 AA.
AC Q03131;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Erythronolide synthase, modules 1 and 2 (EC 2.3.1.94) (ORF 1) (6-
DE deoxyerythronolide B synthase 1) (DEBS 1).
GN ERYA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Actinobacteria; Actinobacteriales (class); Actinobacteridae;
OC Actinomycetales; Pseudonocardiales; Pseudonocardiaceae;
OC Saccharopolyspora.
OX NCBI_TaxID=1836;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9122065; PubMed=2024119;
RA Donadio S., Steaver M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT "Modular organization of genes required for complex polyketide
RT biosynthesis.";
RL Science 252:675-679 (1991).
RN [2]
RP SEQUENCE OF 3474-3491 FROM N.A.
RX MEDLINE=93231529; PubMed=836127;
RA Donadio S., Steaver M.J.;
RT "S1136, an insertion element in the erythromycin gene cluster of
RT Saccharopolyspora erythraea.";
RL Gene 126:147-151 (1993).
CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
CC deoxyerythronolide B.
CC -1- COFACTOR: NADP; CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTHETHEINES.
CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
CC BIOSYNTHESIS.
CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH
CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THIS ERYA SHOWING 3
CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
CC SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS
CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
CC RESPECTIVELY.
CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),
CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACTYL CARRIER
CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
CC OF THE FULL-LENGTH CHAIN.
CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
CC -1- SIMILARITY: CONTAINS 3 ACTYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63676; AAA26493.2; -;
DR EMBL; L07625; AAA26504.1; -;
DR HSSP; P25715; IMLA.
DR InterPro; IPR001218; ADH_short.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Pantane_attach.
DR Pfam; PF00109; adh_short; 1.
DR Pfam; PF00550; pp-binding; 2.
DR Pfam; PF00698; Acyl_transf; 3.
DR Pfam; PF02801; ketoacyl-synt C; 2.

DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
 DR PROSITE; PS00065; B KETOACYL SYNTHASE; 2.
 DR PROSITE; PS00075; ACP DOMAIN; 3.
 KW Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KM Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1972
 FT DOMAIN 1 1979 3491
 FT DOMAIN 1 375
 FT DOMAIN 414 484
 FT DOMAIN 503 961
 FT DOMAIN 1030 1356
 FT DOMAIN 1611 1794
 FT DOMAIN 1888 1958
 FT DOMAIN 1979 2441
 FT DOMAIN 2507 2854
 FT DOMAIN 3055 3237
 FT DOMAIN 3334 3404
 FT ACT_SITE 145 145
 FT BINDING 447 677
 FT ACT_SITE 677 677
 FT BINDING 1128 1128
 FT NP_BIND 1614 1660
 FT BINDING 1921 1921
 FT ACT_SITE 2148 2148
 FT ACT_SITE 2598 2598
 FT NP_BIND 3058 3104
 FT BINDING 3167 3167
 SQ SEQUENCE 3491 AA; 365022 MW; 682BFC32C90FA8C4 CRC64;
 Query Match 49.6%; Score 202; DB 1; Length 3491;
 Best Local Similarity 52.4%; Pred. No. 7e-14;
 Matches 43; Conservative 10; Mismatches 29; Indels 0; Gaps 0;
 QY 1 LAHGTSDQRAALLEVRDHAVALRHADPKAIPDPSFALGSDSLTAVENRLIKAT 60
 DB 1876 LASLPAPERERALELVASHAAVLGHASARVPADQAFALGVDSLSALELRRLGAT 1935
 QY 61 GLRLPSVLVDPHPTPAKLAVAL 82
 DB 1936 GVRLEPTTVFPHDPVRLTAAL 1957
 RESULT 4
 OL56 STRAT STANDARD; PRT: 3519 AA.
 ID OL56 STRAT STANDARD; PRT: 3519 AA.
 AC Q07017; (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Oleandomycin polyketide synthase, modules 5 and 6.
 GN ORF6.
 OS Streptomyces antibioticus.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1890;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150470; PubMed=8107683;
 RA Swan D.G., Rodriguez A.M., Vilches C., Mendez C., Salas J.A.;
 RT "Characterisation of a Streptomyces antibioticus gene encoding a type
 RT I polyketide synthase which has an unusual coding sequence.";
 RL Mol. Gen. Genet. 242:358-362(1994).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE BIOSYNTHESIS OF THE OLEANDOMYCIN
 CC LACTONE RING.
 CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; L09654; AAA19695.1; --
 DR HSSP; P25715; IMLA
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantane-attach.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00550; ketoacyl-synt. 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl_transf; 2.
 DR Pfam; PF00975; Thioesterase; 1.
 DR Pfam; PF02801; ketoacyl-synt C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
 DR PROSITE; PS00065; B KETOACYL SYNTHASE; 2.
 DR PROSITE; PS00075; ACP DOMAIN; 2.
 KW Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;
 KM Phosphopantetheine; Multifunctional enzyme; Repeat.
 FT DOMAIN 1 3519
 FT DOMAIN 32 501
 FT DOMAIN 569 890
 FT DOMAIN 1200 1382
 FT DOMAIN 1487 1561
 FT DOMAIN 1686 2156
 FT DOMAIN 2220 2541
 FT DOMAIN 2856 3038
 FT DOMAIN 3141 3215
 FT ACT_SITE ? ?
 FT BINDING ? ?
 FT DOMAIN 3270 3519
 FT ACT_SITE 210 210
 FT ACT_SITE 660 660
 FT NP_BIND 1203 1249
 FT BINDING 1524 1524
 FT ACT_SITE 1859 1859
 FT ACT_SITE 2311 2311
 FT NP_BIND 2859 2905
 FT BINDING 3178 3178
 SQ SEQUENCE 3519 AA; 368561 MW; 41AE78AAAB61F66 CRC64;
 Query Match 45.7%; Score 186; DB 1; Length 3519;
 Best Local Similarity 50.6%; Pred. No. 4e-12;
 Matches 40; Conservative 12; Mismatches 27; Indels 0; Gaps 0;
 QY 1 LAHGTSDQRAALLEVRDHAVALRHADPKAIPDPSFALGSDSLTAVENRLIKAT 60
 DB 3133 LSMILSPAGREALELVKLVTEAAVLGHSAQDVPAERAFKEIGDSLTAVALNRLAAT 3192
 QY 61 GLRLPSVLVDPHPTPAKLAVAL 79
 DB 3193 GTRLPASAVDPDHPMAALA 3211
 RESULT 5
 PKSL BACSU STANDARD; PRT: 4427 AA.
 ID PKSL BACSU STANDARD; PRT: 4427 AA.
 AC Q05470; (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative polyketide synthase pkel (PKS).
 GN PKSL OR PKSX OR PKSA OR OUTG.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=168 / pkl424;
 RX MEDLINE=93345824; PubMed=8344529;
 RA Scott C., Piatelli M., Cuzzone A., Perani P., Tognoni A., Grandi G.,
 RA Galizzi A., Albertini A.M.;
 RT "A Bacillus subtilis large ORF coding for a polypeptide highly

RT similar to polyketide synthases.",
 RN Gene 130:65-71(1993).
 RL [2]
 RC STRAIN=168;
 RP SEQUENCE FROM N.A.
 RX Kunitz F., Ogasawara N., Moszer I., Albertini A.M., Allison G.,
 RA Aaevedo V., Bertero M.G., Beesler P., Bolotin A., Borshet S.,
 RA Bortles R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brunel C.V., Caldwell B., Capuano N., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Duberhoft A., Enlrich S.D., Emerson P.T.,
 RA Enlian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fitz C., Fujita M., Goffeau A., Gollighly E.J., Grandi G.,
 RA Gilm S.Y., Glaeser P., Goffeau A., Gollighly E.J., Grandi G.,
 RA Gillepeli G., Guy B.J., Haga K., Halech J., Harwood C.R., Henan L.,
 RA Hilbert H., Holappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Jorls B., Karamata D., Kasehara Y., Klier-Bianchard M., Klein C.,
 RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Maeda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Potwolik S., Prescott A.M.,
 RA Pressac E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
 RA Toso V., Uchiyama S., Vandenbol M., Vannier P., Vassalotti A.,
 RA Viari A., Wamburt R., Wedler E., Wedler H., Wetzengesser T.,
 RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumelein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.",
 RT Nature 390:249-256(1997).
 RL [3]
 RN SEQUENCE OF 3619-4427 FROM N.A.
 RC STRAIN=168 / PB1424;
 RL Grandi G.,
 RL Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR
 CC THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN
 CC SECONDARY METABOLISM.
 CC -1- COFACTOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES
 CC (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS 5 ACYL CARRIER DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z14098; CAAT78479.1; -
 CC EMBL: U11039; AAA65145.1; -
 CC EMBL: Z59113; CAB1602.1; -
 CC EMBL: Z51133; CAAB4504.1; -
 CC PIR: S25021; S25021.
 CC Subtilase; BG10698; PKSL.
 CC InterPro: IPR002198; ADH_short.
 CC InterPro: IPR000794; ketoacyl-synt.
 CC InterPro: IPR003880; Ppanne_attach.
 CC Pfam: PF00106; adh_short; 1.
 CC Pfam: PF00109; ketoacyl-synt; 4.
 CC Pfam: PF00550; pp-binding; 5.
 CC Pfam: PF02801; ketoacyl-synt C; 4.
 CC PROSITE: PS00012; PHOSPHOPANTETHEINE; 5.
 CC PROSITE: PS00066; B_KETOACYL_SYNTHASE; 1.
 CC PROSITE: PSS0075; ACP_DOMAIN; 5.

KW Transferase: Acyltransferase; Antibiotic biosynthesis; NADP;
 KM Phosphopantetheine; Multifunctional enzyme; Repeat; Complete proteome.
 FT DOMAIN 211 280
 FT DOMAIN 382 759
 FT DOMAIN 937 1115
 FT DOMAIN 1409 1602
 FT DOMAIN 1687 1759
 FT DOMAIN 1876 2253
 FT DOMAIN 2491 2560
 FT DOMAIN 2632 2701
 FT DOMAIN 2823 3182
 FT DOMAIN 3575 3776
 FT DOMAIN 3854 3923
 FT DOMAIN 4019 4373
 FT BINDING 243 243
 FT BINDING 1723 1723
 FT BINDING 2523 2523
 FT BINDING 2664 2664
 FT BINDING 3886 3886
 SQ SEQUENCE 4427 AA; 493398 MW; 9612521E561AB9F2 CRC64;
 Query Match 25.7%; Score 104.5; DB 1; Length 4427;
 Best Local Similarity 34.2%; Pred. No. 0.0045;
 Matches 27; Conservative 16; Mismatches 35; Indels 1; Gaps 1;
 Qy 4 QTSADPRAALLBTRDVAVALRHADPKAIAPDSFRAGPSLTVFRNLIKATGLR 63
 Db 2482 EFTTIRAEALCDELDTGLAEVL-YMDQNEIDPEARIDGMSITLWIKAIKQYGS 2540
 Qy 64 LPVSLVDHPTPAKAVHL 82
 Db 2541 LNTKTYDYPTTRDFAVYL 2559
 RESULT 6
 PPSE_MYCTU STANDARD; PRT; 1538 AA.
 ID PPSE_MYCTU
 AC Q10978; O53234;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phenolphthalein synthesis polyketide synthase ppse.
 GN PPSE OR RV2932 OR MT3002 OR MTCY338.21 OR MT0011.01.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OK NCBI_TaxId=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigler K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whittam S., Whittam B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.",
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Mayhew L.A., Esmailova M.D., Salzberg S.L.,
 RA Delcher A., Utermack T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.,
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.",
 RT Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLPTHIOCEROL
CC SYNTHESIS.
CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETHEINE (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AL021070; CAA15929.1; -
DR EMBL; AE007122; AAK47329.1; -
DR TIGR; MT3002; -
DR Tuberculist; RV2932; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Pantane_attach.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF00550; pp-binding; 1.
DR Pfam; PF00698; Acyl_transf; 1.
DR Pfam; PF02801; ketoacyl-synt C; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; FALSE_NEG.
DR PROSITE; PS50075; ACP_DOMAIN; 1.
DR Hypothetical protein; Multifunctional enzyme; Oxidoreductase;
DR Transferase; NADP; Phosphopantetheine; Complete proteome.
FT DOMAIN 1425 1495 ACP_CARRIER (ACP).
FT BINDING 1458 1458 PHOSPHOPANTHETHEINE (POTENTIAL).
SQ SEQUENCE 1538 AA; 162527 MW; B55E2A042AD00CC CRC64;

Query Match 24.1%; Score 98; DB 1; Length 1538;
Best Local Similarity 34.7%; Pred. No. 0.0075;
Matches 25; Conservative 13; Mismatches 30; Indels 4; Gaps 1;

QY 15 ELVVDH-----AVLVRHADPKAIPDQSFALGPDSLTAVFRRLLIKATGRLRPVSLVF 70
DB 1423 DMLPDHGAALATVGMPTPEPLDPSAGFOLGMDLSMTVLQRLSESLGEFLPASVVF 1482
QY 71 DHPTPAKLAHL 82
DB 1483 DPTVTVSLTDTL 1494

RESULT 7
PSSA MYCTU STANDARD; PRT; 1876 AA.
ID PSSA MYCTU
AC 010977;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phenolphthiocerol synthesis polyketide synthase ppsa.
GN PSSA OR RV2931 OR MT3000 OR MTCY338.20.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteriae; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gae S., Barry C.E. III, Tekaia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., Mclean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the

RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleishmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
RA Bishai W., Ueberback T., Weldman J., Khouri H., Gill J., Mikula A.,
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLPTHIOCEROL
CC SYNTHESIS.
CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETHEINES (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; Z74697; CAA98988.1; -
DR EMBL; AE007122; AAK47328.1; -
DR TIGR; MT3000; -
DR Tuberculist; RV2931; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Pantane_attach.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF00550; pp-binding; 2.
DR Pfam; PF00698; Acyl_transf; 1.
DR Pfam; PF02801; ketoacyl-synt C; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR PROSITE; PS50075; ACP_DOMAIN; 2.
DR Multifunctional enzyme; Oxidoreductase; Transferase; NADP; Repeat;
KW Phosphopantetheine; Complete proteome.
FT DOMAIN 7 80 ACP_CARRIER (ACP) 1.
FT NP_BIND 1764 1833 ACP_CARRIER (ACP) 2.
FT BINDING 62 62 NADP (POTENTIAL).
FT ACT_SITE 273 273 PHOSPHOPANTHETHEINE (POTENTIAL).
FT ACT_SITE 720 720 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT BINDING 1796 1796 MALONYLTRANSFERASE (BY SIMILARITY).
FT BINDING 624 624 PHOSPHOPANTHETHEINE (POTENTIAL).
FT CONFLICT 877 877 D -> E (IN REF. 2).
FT CONFLICT 877 877 R -> H (IN REF. 2).
FT CONFLICT 1323 1323 G -> S (IN REF. 2).
SQ SEQUENCE 1876 AA; 198834 MW; D9783DBD48792110 CRC64;

Query Match 23.2%; Score 94.5; DB 1; Length 1876;
Best Local Similarity 32.5%; Pred. No. 0.023;
Matches 25; Conservative 17; Mismatches 34; Indels 1; Gaps 1;

QY 6 SADORALALTELVDHVAVALVRHADPKAIPDQSFALGPDSLTAVFRRLLIKATGRLRP 65
DB 1757 ATEVRSELBEGGLRIIAAEIR-VPEKELDTPDFALGSLNSLWAMVIRREABQFVGIELS 1815
QY 66 VSLVDFHPTPAKLAHL 82
DB 1816 ATMLFNHPVTKSLASVYL 1832

RESULT 8
DHBF BACSU STANDARD; PRT; 1278 AA.
ID DHBF BACSU
AC P45745;
DT 01-NOV-1995 (Rel. 32, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable serine activating enzyme.
 GN DHPF.
 OS *Bacillus subtilis*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunze F., Ogatawara N., Moszer I., Albertini A.M., Alloni G.,
 Aaevedo V., Bertero M.G., Beselares P., Bolotin A., Borchet S.,
 Borriera R., Boursier L., Brans A., Braun M., Briegleb S.C., Bron S.,
 Brouillet S., Brunel C.V., Caldwell B., Capuano V., Carter N.M.,
 Choi S.K., Codari J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 Denize F., Devine K.M., Duesterhoft A., Ehlich S.D., Emmerson P.T.,
 Enlian K.D., Ertugron J., Fadret C., Ferrari E., Fougere D.,
 Fritz S.Y., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 Galtier G., Gasser P., Goffeau A., Golightly E.J., Grandi G.,
 Guileppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 Hilbert H., Holappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 Joris B., Karamata D., Kasahara Y., Klaerr-Binchard M., Klein C.,
 Kobayashi Y., Koetter P., Koningsstein G., Krog S., Kumano M.,
 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,
 Parro V., Pohl T.M., Portetelle D., Portetlik S., Prescott A.M.,
 Priesen E., Pujic P., Purnelle B., Repoport G., Rey M., Reynolds S.,
 Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadleir Y.,
 Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
 Totsu V., Uchiyama S., Vandenbol M., Vannier P., Vassatoti A.,
 Varti A., Wambut R., Wedler E., Wedler H., Weitzneger T.,
 Winters P., Wipac A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Dandhin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*,"
 RT Nature 390:249-256(1997).
 RL [2]
 RN SEQUENCE OF 1-95 FROM N.A.
 RP STRAIN=168 / Marburg;
 RC MEDLINE=8616538; PubMed=8550523;
 RX Rowland B.M., Taber H.M.;
 RA "Duplicate isochlorismate synthase genes of *Bacillus subtilis*:
 RT regulation and involvement in the biosyntheses of menaquinone and
 RT 2,3-dihydroxybenzoate,"
 RT J. Bacteriol. 178:854-861(1996).
 CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETINE (BY
 CC similarity)
 CC -1- PATHWAY: 2,3-dihydroxybenzoate biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL; Z99120; CAB15186.1; -
 CC EMBL; U26444; AAC44634.1; -
 CC HSSP; P14687; 1AMU.
 CC Subtilin; BG1243; dbp.
 CC InterPro; IPR000873; AMP-bind.
 CC InterPro; IPR001242; Condensatn.
 CC InterPro; IPR003880; Pantene_attach.

DR Pfam; PF00501; AMP-binding; 1.
 DR Pfam; PF00550; PP-binding; 1.
 DR Pfam; PF00668; Condensation; 2.
 DR PRINTS; PR00154; AMPBINDING.
 DR PROSITE; PS00012; PHOSPHOPANTHETINE; FALSE_NEG.
 DR PROSITE; PS00455; AMP-BINDING; 1.
 DR PROSITE; PS50075; ACP_DOMAIN; 1.
 KM Lysase; Phosphopantetheine; Complete proteome.
 FT DOMAIN 965 1033 ACYL CARRIER (ACP).
 FT BINDING 965 996 PHOSPHOPANTHETINE (POTENTIAL).
 SQ SEQUENCE 1278 AA; 142068 MW; B44FBFF3FCB085B4 CRC64;
 Query Match 21.4%; Score 87; DB 1; Length 1278;
 Best Local Similarity 35.8%; Pred. No. 0.099;
 Matches 29; Conservative 9; Mismatches 37; Indels 6; Gaps 2;
 Db 5 TSADQRAALL--ELVRDHAVALRRAADPKALPPDSFRLGDSLTAEFRLLKATG 61
 955 TSVSDRAAPRPPQREILCDLFAEVLGLA--RVGIDDSFELGSHSLAARLMSRIRVWG 1011
 Db 62 LRLVSLVDFDPRAKLAVAL 82
 1012 AELGIKALFDEPTVAGLAHL 1032
 RESULT 9
 MSAS PENPA STANDARD; PRT; 1774 AA.
 ID MSAS PENPA
 AC P22367;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 6-methylsalicylic acid synthase (EC 2.3.1.-) (MSAS).
 OS *Penicillium patulum* (penicillium griseofulvum).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; *Penicillium*.
 OX NCBI_TaxID=5078;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=DSM 62862;
 RX MEDLINE=9106137; PubMed=2209605;
 RA Beck J., Ripka S., Slegner A., Schiltz E., Schweizer E.;
 RT "The multifunctional 6-methylsalicylic acid synthase gene of
 RT *Penicillium patulum*. Its gene structure relative to that of other
 RT polyketide synthases,"
 RT Eur. J. Biochem. 192:487-498(1990).
 CC -1- FUNCTION: This multifunctional enzyme is a polyketide synthase.
 CC It catalyzes a total of 11 steps by seven different component
 CC enzymes, in the biosynthesis of the antibiotic patulin.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + 3 malonyl-CoA + NADPH =
 CC 6-methylsalicylic acid + NADP(+) + 3 CoA + 3 CO(2) + H(2)O.
 CC -1- PATHWAY: Patulin biosynthesis.
 CC -1- SUBUNIT: HOMOMULTIMER.
 CC -1- INDUCTION: IN THE LATE LOGARITHMIC GROWTH PHASE.
 CC -1- SIMILARITY: WITH MOST POLYKETIDE SYNTHASES, SPECIALLY WITH RAT
 CC FATTY ACID SYNTHASE, AND WITH OTHER ENZYMES SUCH AS LIPASES AND
 CC THIOLEASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL; X55776; CAA39295.1; -
 CC PIR; S13178; S13178.
 CC InterPro; IPR001227; Ac transferase.
 CC InterPro; IPR000794; Ketocacyl-synt.
 CC InterPro; IPR003880; Ketocacyl_attach.
 CC Pfam; PF00109; ketocacyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 1.

DR Pfam; PF00698; Acyl transferase; 1.
 DR Pfam; PF02801; ketoacyl-synt C; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE; PS50075; ACP_DOMAIN; 1.
 DR Multifunctional enzyme; Oxidoreductase; Antibiotic biosynthesis;
 KM Transferase; NADP; Phosphopantetheine.
 FT DOMAIN 186 238 ACYLTRANSFERASE (AT).
 FT DOMAIN 642 676 ACETYL/MALONYL TRANSFERASES.
 FT DOMAIN 1403 1450 2-OXOACYL REDUCTASE.
 FT DOMAIN 1700 1769 ACYL CARRIER (ACP).
 FT NP_BIND 1419 1424 NADP (POTENTIAL).
 FT ACT_SITE 204 204 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT_SITE 653 653 MALONYLTRANSFERASE (BY SIMILARITY).
 FT BINDING 1732 1732 PHOSPHOPANTETHEINE (BY SIMILARITY).
 SQ SEQUENCE 1774 AA; 190732 MW; 05ED5DD10863f938 CRC64;

Query Match 20.3%; Score 82.5; DB 1; Length 1774;
 Best Local Similarity 36.0%; Pred. No. 0.44;
 Matches 27; Conservative 9; Mismatches 30; Indels 9; Gaps 2;

DR 10 RAALLEVRDHVAANVR---HADPKAIAPDQSFALGFDLSLAVERRNLLIKATGRLP 65
 Db 1697 KAYLDEKIRGCVAKVLTQMTAEVDVSKAALAD---LGVDSVMTVTLRQQLTLTKIAPV 1751

QY 66 VSLVFDHPTAKLAV 80
 Db 1752 PTLTWSHPTVSHLAV 1766

RESULT 10
 MCAS MYCBO STANDARD; PRT; 2110 AA.
 AC 002251;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, last sequence update)
 DT 15-JUL-1998 (Rel. 36, last annotation update)
 DE Mycocerosic acid synthase.
 GN MAS.
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 ON NCBI_TaxID=1765;
 RX MEDLINE=92406887; PubMed=1527058;
 RA Metchur M., Kolatukudy P.E.;
 RT "Molecular cloning and sequencing of the gene for mycocerosic acid
 synthase, a novel fatty acid elongating multifunctional enzyme, from
 Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guerin.";
 RT J. Biol. Chem. 267:19388-19395 (1992).
 CC -1- FUNCTION: CATALYZES THE ELONGATION OF N-PATTY AYL-COA WITH
 METHYLMALONYL-COA (NOT MALONYL-COA) AS THE ELONGATING AGENT TO
 FORM MYCOCEROXYL LIPIDS.
 CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTETHEINE.
 CC -1- SUBUNIT: HOMODIMER WHOSE MONOMERS PROBABLY HAVE A HEAD TO TAIL
 ARRANGEMENT.
 CC -1- SUBCELLULAR LOCATION: Membrane-associated.
 CC -1- SIMILARITY: CONTAINS 1 AYL CARRIER DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M95908; AAA25369.1; -.
 DR FTR; B44110; B44110.
 DR HSP; P73283; IBSM.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh_zn_family.

DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Pantane_attach.
 DR Pfam; PF00109; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR Pfam; PF00698; Acyl transferase; 1.
 DR Pfam; PF02801; ketoacyl-synt C; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE; PS50075; ACP_DOMAIN; 1.
 DR Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 KM Transferase; Hydrolase; Oxidoreductase; Ligase; NADP; Membrane.
 FT DOMAIN 1 430 BETA-KETOACYL SYNTHASE.
 FT DOMAIN 533 852 ACYL TRANSFERASE.
 FT DOMAIN ? ? ENOYL REDUCTASE.
 FT DOMAIN ? ? BETA-KETOACYL REDUCTASE.
 FT DOMAIN 2026 2096 ACYL CARRIER (ACP).
 FT ACT_SITE 177 177 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT_SITE 623 623 ACYL TRANSFERASES (BY SIMILARITY).
 FT NP_BIND 1561 1578 NADP (ER).
 FT NP_BIND 1765 1780 NADP (KR).
 FT BINDING 2059 2059 PHOSPHOPANTETHEINE (BY SIMILARITY).
 SQ SEQUENCE 2110 AA; 225577 MW; C65821507155300 CRC64;

Query Match 18.9%; Score 77; DB 1; Length 2110;
 Best Local Similarity 31.4%; Pred. No. 2.1;
 Matches 22; Conservative 13; Mismatches 31; Indels 4; Gaps 1;

QY 13 LLEVRDHVAANVRHADPKAIAPDQSFALGFDLSLAVERRNLLIKATGRLPVSIVPDH 72
 Db 2030 LRRLIVEQASVILR---RTIDADRSEFIEYGLSLQMLEWRTHTVETETGIRLPKVIATN 2085

QY 73 PTPAKLAVHL 82
 Db 2086 KTRALAQYL 2095

RESULT 11
 LYS2 YEAST STANDARD; PRT; 1392 AA.
 ID LYS2 YEAST
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-MAR-1992 (Rel. 21, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Aminoadipate-semialdehyde dehydrogenase large subunit (EC 1.2.1.31)
 DE (Alpha-aminoadipate reductase) (Alpha-AR).
 GN LYS2 OR YBR115C OR YER0910.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 ON NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=91192607; PubMed=2013406;
 RA Morris M.E., Jinks-Robertson S.;
 RT "Nucleotide sequence of the LYS2 gene of Saccharomyces cerevisiae:
 RT homology to Bacillus brevis tyrocidine synthetase 1.";
 RT Gene 98:141-145 (1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=S288c;
 RC MEDLINE=95208357; PubMed=7900426;
 RA Mannhaupt G., Stucka R., Ehme S., Vetter I., Feldmann H.;
 RT "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
 RT (Yeast 10:1363-1381 (1994)).
 RN [3]
 RP SEQUENCE OF 1-150 AND 1209-1392 FROM N.A.
 RP MEDLINE=87106859; PubMed=3542721;
 RA Fleig U.N., Pridmore R.D., Philippen P.;
 RT "Construction of LYS2 catridges for use in genetic manipulations of
 RT Saccharomyces cerevisiae.";
 RT Gene 46:237-245 (1986).
 RN [4]

RP SEQUENCE OF 1083-1392 FROM N.A.
 RC STRAIN=8288c;
 RX MEDLINE=92327848; PubMed=1626431;
 RA Mannhaupt G., Stucke R., Ehme S., Vetter I., Feldmann H.;
 RT "Molecular analysis of yeast chromosome II between CND1 and LYS2: the
 RT excision repair gene RAD16 located in this region belongs to a novel
 RT group of double-finger proteins.";
 RL Yeast 8:397-408(1992).
 RN [5]
 RP SEQUENCE OF 1-130 FROM N.A.
 RC STRAIN=8288c;
 RX MEDLINE=94039074; PubMed=7916691;
 RA Schaeff-Geretschläger I., Mannhaupt G., Vetter I., Zimmermann F.K.,
 RA Feldmann H.;
 RT "FKL2, a second transketolase gene of Saccharomyces cerevisiae.
 RT Cloning, sequence and deletion analysis of the gene.";
 RL Eur. J. Biochem. 217:487-492(1993).
 CC -1- FUNCTION: CATALYZES THE ACTIVATION OF ALPHA-AMINOADIPATE BY
 CC ALPHA-AMINOADIPATE AND THE REDUCTION OF ACTIVATED
 CC ALPHA-AMINOADIPATE BY NADPH.
 CC -1- CATALYTIC ACTIVITY: L-2-aminoadipate 6-semialdehyde + NADP(+) +
 CC H(2)O = L-2-aminoadipate + NADPH.
 CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTETHEINE
 CC (POTENTIAL).
 CC -1- PATHWAY: Lysine biosynthesis, sixth step.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M36287; AAA34747.1; -
 DR EMBL: X66247; CAA46975.1; -
 DR EMBL: X78993; CAA55617.1; -
 DR EMBL: X235984; CAA85072.1; -
 DR EMBL: X73532; CAA51938.1; -
 DR PIR: J00448; YGBYAD.
 DR HSSP: P14687; 1AMU.
 DR SGD: S0000319; LYS2.
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR003880; Pantene_attach.
 DR Pfam: PF00501; AMP-binding; 1.
 DR Pfam: PF00550; pp-binding; 1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE: PS00455; AMP-BINDING; 1.
 DR PROSITE: PS50075; ACP DOMAIN; 1.
 KM Lysine biosynthesis; Oxidoreductase; NADP; Phosphopantetheine.
 FT DOMAIN 848 917 ACYL CARRIER (ACP).
 FT BINDING 880 880 PHOSPHOPANTETHEINE (POTENTIAL).
 SO SEQUENCE 1392 AA; 155345 MW; F0083A80B6C7FB5 CRC64;

Query Match 18.6%; Score 75.5; DB 1; Length 1392;
 Best Local Similarity 31.8%; Pred. No. 2;
 Matches 21; Conservative 12; Mismatches 32; Indels 1; Gaps 1;

QY 17 VRDVAIVLHNDPKALAPDSITAVEPNNLLKATGLLPVSLVVDHTPTA 76
 DB 852 VRDVAIVLHNDPKALAPDSITAVEPNNLLKATGLLPVSLVVDHTPTA 76
 QY 77 KLAIVL 82
 DB 911 AFAAEI 916

RESULT 12
 FAS_RAT

ID FAS_RAT STANDARD; PRT; 2505 AA.
 AC P12785; O64717; O09187; O09190;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Fatty acid synthase [EC 2.3.1.85] (includes: EC 2.3.1.38; EC 2.3.1.39;
 DE EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14).
 GN FASN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89240686; PubMed=2717611;
 RA Amy C.M., Witkowski A., Naggert J., Williams B., Randhawa Z.,
 RA Smith S.;
 RT "Molecular cloning and sequencing of cDNAs encoding the entire rat
 RT fatty acid synthase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3114-3118(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=93075999; PubMed=1339331;
 RA Beck K.F., Schlegelmann R., Stathopoulos I., Klein H., Hoch J.,
 RA Schweitzer M.;
 RT "The fatty acid synthase (FAS) gene and its promoter in Rattus
 RT norvegicus.";
 RL DNA Seq. 2:359-386(1992).
 RN [3]
 RP SEQUENCE OF 75-2505 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Mammary gland;
 RX MEDLINE=89128431; PubMed=2915923;
 RA Schweitzer M., Takabayashi K., Beck K.F., Schlegelmann R.;
 RA "Rat mammary gland fatty acid synthase: localization of the
 RT consistent domains and two functional polyadenylation/termination
 RT signals in the cDNA.";
 RL Nucleic Acids Res. 17:567-586(1989).
 RN [4]
 RP SEQUENCE OF 2085-2505 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=88087240; PubMed=2891707;
 RA Naggert J., Witkowski A., Mikkelson J., Smith S.;
 RA "Molecular cloning and sequencing of a cDNA encoding the thioesterase
 RT domain of the rat fatty acid synthetase.";
 RL J. Biol. Chem. 263:1146-1150(1988).
 RN [5]
 RP SEQUENCE OF 1921-2324 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=87246646; PubMed=3109907;
 RA Witkowski A., Naggert J., Mikkelson J., Smith S.;
 RT "Molecular cloning and sequencing of a cDNA encoding the acyl carrier
 RT protein and its flanking domains in the mammalian fatty acid
 RT synthetase.";
 RL Eur. J. Biochem. 165:601-606(1987).
 CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA = a
 CC long-chain fatty acid + (N+1) COA + 2N NADP(+).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl]-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxypanthoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-

2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-carrier protein] + oleate.
 -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC between the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 DR EMBL; M76767; AAA4679.1; -;
 DR EMBL; X62888; CAA44679.1; -;
 DR EMBL; X62889; CAA44680.1; -;
 DR EMBL; X13415; CAA31780.1; -;
 DR EMBL; X13527; CAA31882.1; -;
 DR EMBL; J03514; AAA41144.1; -;
 DR PIR; A30313; XRTFA.

DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Acln_zn_family.
 DR InterPro; IPR007994; Ketoacyl-synt.
 DR InterPro; IPR003880; Ppantne_atlch.
 DR InterPro; IPR000051; SAM_bind.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR Pfam; PF00698; acyl_transf; 1.
 DR Pfam; PF02801; thioesterase; 1.
 DR Pfam; PF02801; ketoacyl-synt; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE; PS50075; ACP_DOMAIN; 1.
 KM Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 KM Hydroxylase; Oxidoreductase; Transferase; Lyase; NADP;
 KM Pyridoxal phosphate.

FT DOMAIN 1 413 BETA-KETOACYL SYNTHASE.
 FT DOMAIN 429 817 ACYL AND MALONYL TRANSFERASES.
 FT DOMAIN 1629 1857 ENOYL REDUCTASE.
 FT DOMAIN 1858 2113 BETA-KETOACYL REDUCTASE.
 FT DOMAIN 2118 2174 ACYL CARRIER (ACP).
 FT DOMAIN 2202 2505 THIOESTERASE.
 FT ACT_SITE 161 161 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT_SITE 581 581 MALONYLTRANSFERASE (BY SIMILARITY).
 FT NP_BIND 1665 1682 NADP (ER).
 FT BINDING 1698 1698 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT NP_BIND 1765 1780 NADP (KR).
 FT BINDING 2151 2151 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT ACT_SITE 2302 2302 THIOESTERASE (BY SIMILARITY).
 FT ACT_SITE 2475 2475 THIOESTERASE (BY SIMILARITY).
 FT ACT_SITE 878 878 BETA-HYDROXYACYL DEHYDRATASE (BY SIMILARITY).
 FT CONFLICT 871 871 S -> P (IN REF. 3).
 FT CONFLICT 1967 1968 MV -> IL (IN REF. 5).
 FT CONFLICT 2085 2085 C -> P (IN REF. 4).
 FT CONFLICT 2106 2106 A -> V (IN REF. 1 AND 5).
 FT CONFLICT 2295 2295 Y -> H (IN REF. 1 AND 5).
 SQ SEQUENCE 2505 AA; 272647 MW; 5810EC13D57F3114 CRC64;

Query Match 17.6%; Score 71.5; DB 1; Length 2505;
 Best Local Similarity 37.9%; Pred. No. 10;
 Matches 25; Conservative 7; Mismatches 29; Indels 5; Gaps 2;

QY 1 LAHGTSDGQALAEIVRDHVAALVLRADPKALAPDSFRALGDSITLAVEFRNLITKAT 60
 DB 2111 VAHGDGEAQR-----DLVK-AVAHILGIRDLAGINLDSLADLGDSIMGEVQRLEH 2165

QY 61 GLRLPV 66
 DB 2166 DLVLPI 2171

RESULT 13
 ACP_STRCO
 ID_ACP_STRCO STANDARD; PRT; 92 AA.

AC 09RK61;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ACP1 carrier protein (ACP).
 GN ACP OR SC00549 OR SCF11.29C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID:1902;
 [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neil S.,
 RA Rabinowitz S., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrrell B.G., Parkhill J.,
 RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)."
 RL Nature 417:141-147(2002).

CC -1- FUNCTION: Carrier of the growing fatty acid chain in fatty acid
 biosynthesis (By similarity).
 CC -1- PATHWAY: De novo fatty acid biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- PTM: 4'-phosphopantetheine is transferred from CoA to a specific
 serine of apo-ACP by acps. This modification is essential for
 activity because fatty acids are bound in thioester linkage to the
 sulfhydryl of the prosthetic group (By similarity).
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC between the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AL132662; CAB59606.1; -;
 DR InterPro; IPR003880; Ppantne_atlch.
 DR Pfam; PF00550; pp-binding; 1.
 DR PROSITE; PS50075; ACP_DOMAIN; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; FALSE NEG.
 KM Lipid synthesis; Fatty acid biosynthesis; Phosphopantetheine;
 KM Complete proteome.

KW BINDING 44 44 PHOSPHOPANTETHEINE (BY SIMILARITY).
 FT BINDING 44 44
 SQ SEQUENCE 92 AA; 9876 MW; 8A1800D2395AA2B3 CRC64;

Query Match 17.4%; Score 71; DB 1; Length 92;
 Best Local Similarity 30.4%; Pred. No. 0.33;
 Matches 24; Conservative 15; Mismatches 38; Indels 2; Gaps 2;

QY 5 TSADGQALAEIVRDHVAALVLRADPKALAPDSFRALGDSITLAVEFRNLITKATGR 63
 DB 3 STDER-QULDEIRDLAATVEDLTVEEIGPSSLODDLGVSILARLVAALIEDRWQIE 61

QY 64 LPVSLVDFHPTPAKLAHL 82
 DB 62 VPQEQADRLLTVRQIAHL 80

RESULT 14
 HTSI_COCCA

RT of srfaA of the Bacillus subtilis chromosome.",
 RL Microbiology 141:277-279(1995).
 CC -1- FUNCTION: THIS PROTEIN IS A MULTIFUNCTIONAL ENZYME ABLE TO
 CC ACTIVATE AND POLYMERIZE THE AMINO ACIDS LEU, GLU, ASP AND VAL.
 CC ACTIVATION SITES FOR THESE AA CONSIST OF INDIVIDUAL DOMAINS.
 CC -1- COFACTOR: CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTETHEINES.
 CC -1- PATHWAY: Cyclic peptide antibiotic surfactin biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D13262; BAA02522.1; -;
 CC EMBL: X70356; CAA49816.1; -;
 CC EMBL: D50453; BAA08982.1; -;
 CC EMBL: M59939; AAA22815.1; -;
 CC EMBL: M64702; AAA22816.1; -;
 CC EMBL: D30762; BAA21034.1; -;
 CC EMBL: Z99105; CAB12142.1; -;
 CC PIR: S35517; S35517.
 CC HSSP: P14687; IAMU.
 DR Sublistet; BG10168; srfaA.
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR001242; Condensatn.
 DR InterPro: IPR003880; Pantene attach.
 DR Pfam: PF00501; AMP-binding; 3.
 DR Pfam: PF00550; pp-binding; 3.
 DR Pfam: PF00668; Condensation; 4.
 DR PRINTS; PR00154; AMPBINDING.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
 DR PROSITE; PS00455; AMP BINDING; 3.
 DR PROSITE; PS50075; ACP DOMAIN; 3.
 DR Ligase; Antibiotic biosynthesis; Phosphopantetheine; Sporulation;
 KW Multifunctional enzyme; Repeat; Complete proteome.
 FT REPEAT ? 1047 DOMAIN 1 (GLU-ACTIVATING).
 FT REPEAT ? 2086 DOMAIN 2 (LEU-ACTIVATING).
 FT REPEAT ? 3114 DOMAIN 3 (LEU-ACTIVATING).
 FT DOMAIN 976 1043 ACYL CARRIER (ACP) 1.
 FT DOMAIN 2015 2082 ACYL CARRIER (ACP) 2.
 FT DOMAIN 3043 3109 ACYL CARRIER (ACP) 3.
 FT BINDING 1006 1006 PHOSPHOPANTETHEINE (POTENTIAL).
 FT BINDING 2045 2045 PHOSPHOPANTETHEINE (POTENTIAL).
 FT BINDING 3073 3073 PHOSPHOPANTETHEINE (POTENTIAL).
 FT BINDING 146 146 I -> M (IN REF. 2 AND 3).
 FT BINDING 151 151 L -> V (IN REF. 2 AND 3).
 FT BINDING 165 165 A -> G (IN REF. 2, 3 AND 4).
 FT BINDING 281 281 T -> Q (IN REF. 2 AND 3).
 FT BINDING 460 460 D -> T (IN REF. 3 AND 3).
 FT BINDING 540 540 P -> A (IN REF. 2 AND 3).
 FT BINDING 562 562 I -> Y (IN REF. 2 AND 3).
 FT BINDING 639 640 GS -> PT (IN REF. 2 AND 3).
 FT BINDING 644 644 R -> P (IN REF. 2 AND 3).
 FT BINDING 647 647 L -> F (IN REF. 2 AND 3).
 FT BINDING 649 649 A -> P (IN REF. 2 AND 3).
 FT BINDING 1026 1026 L -> Q (IN REF. 2 AND 3).
 FT BINDING 1065 1116 SLAQQTIVISQPEDAGVGMPPAALLEGDLIOKLERAF
 FT BINDING 1065 1116 OGILRHESLR -> HHSSSEHTSSASSRMRESHHSN
 FT BINDING 1065 1116 SEGRYSKAGARISGINPTREVE (IN REF. 2 AND
 FT BINDING 1065 1116 3).
 FT BINDING 1132 1133 DS -> VC (IN REF. 2 AND 3).
 FT BINDING 1164 1164 R -> A (IN REF. 2 AND 3).
 FT BINDING 1458 1458 V -> L (IN REF. 2 AND 3).
 FT BINDING 1850 1850 S -> T (IN REF. 2 AND 3).
 FT BINDING 1894 1894 D -> T (IN REF. 2 AND 3).
 FT BINDING 1911 1911 R -> A (IN REF. 2 AND 3).
 FT BINDING 1973 1980 LRCLSKWT -> PAVPTQMD (IN REF. 2 AND 3).

FT CONFLICT 2052 2052 L -> P (IN REF. 2 AND 3).
 FT CONFLICT 2218 2218 H -> Q (IN REF. 2 AND 3).
 FT CONFLICT 2265 2266 QQ -> HR (IN REF. 2 AND 3).
 FT CONFLICT 2291 2291 V -> L (IN REF. 2 AND 3).
 FT CONFLICT 2349 2349 S -> C (IN REF. 2 AND 3).
 FT CONFLICT 2428 2428 R -> S (IN REF. 2 AND 3).
 FT CONFLICT 2611 2612 RC -> AV (IN REF. 2 AND 3).
 FT CONFLICT 2685 2685 E -> ENPE (IN REF. 2 AND 3).
 FT CONFLICT 2755 2755 T -> S (IN REF. 2 AND 3).
 FT CONFLICT 2897 2897 SP -> TA (IN REF. 2 AND 3).
 FT CONFLICT 3025 3025 P -> N (IN REF. 2 AND 3).
 FT CONFLICT 3096 3096 F -> N (IN REF. 2 AND 3).
 FT CONFLICT 3271 3271 A -> S (IN REF. 2 AND 3).
 FT CONFLICT 3316 3316 R -> S (IN REF. 2 AND 3).
 FT CONFLICT 3451 3451 Y -> S (IN REF. 2 AND 3).
 FT CONFLICT 3483 3484 DE -> HQ (IN REF. 2 AND 3).
 FT CONFLICT 3487 3490 DAGL -> HPPF (IN REF. 2 AND 3).
 FT CONFLICT 3493 3493 R -> T (IN REF. 2 AND 3).
 FT CONFLICT 3495 3495 E -> H (IN REF. 2 AND 3).
 FT CONFLICT 3499 3500 GQ -> PH (IN REF. 2 AND 3).
 FT CONFLICT 3507 3507 R -> E (IN REF. 2 AND 3).
 SQ SEQUENCE 3587 AA; 402424 MW; A1E2DABF93JED3A CRC64;

Query Match 17.2%; Score 70; DB 1; Length 3587;
 Best local Similarity 35.7%; Pred. No. 22;
 Matches 15; Conservative 8; Mismatches 19; Indels 0; Gaps 0;

QY 38 SFRALGDSLTAVFNNLLKATGLRPLVSLVEDHPTPKLA 79
 DB 998 NPFETGSHSLKAWTLTKIKETGISRIPLOPLEHPTTALA 1039

Search completed: June 17, 2003, 13:02:36
 Job time : 1.89424 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:55:32 ; Search time 1.83105 Seconds

(without alignments)
4305.183 Million cell updates/sec

Title: US-09-914-286-3_COPY_1979_2060

Perfect score: 407

Sequence: 1 LAHQTSADQRAALLEVRDH.....RLPVSIVFDHPTPAKLAVHL 82

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	229	56.3	3739	2 T17410	polyketide synthas
2	225	55.3	1562	2 T17411	polyketide synthas
3	224	55.0	1763	2 T17465	rifamycin polyketi
4	223	54.8	2100	2 T03223	probable polyketid
5	221	54.3	5069	2 T17464	rifamycin polyketi
6	221	54.3	6420	2 T30283	polyketide synthas
7	216.5	53.2	1762	2 T03232	probable polyketid
8	216	53.1	3413	2 T17467	rifamycin polyketi
9	216	53.1	4151	2 G70944	probable polyketid
10	215	52.8	8563	2 T30226	polyketide synthas
11	214.5	52.7	10223	2 T30225	polyketide synthas
12	212	52.1	1728	2 T17466	rifamycin polyketi
13	212	52.1	6260	2 T30228	polyketide synthas
14	205	50.4	3172	2 S22012	erythronolide synt
15	205	50.4	3178	2 S13595	6-deoxyerythronoli
16	203	49.9	1937	2 T03224	probable polyketid
17	202.5	49.8	7576	2 T17428	FK506 polyketide s
18	202	49.6	4391	2 T43231	erythronolide synt
19	199	48.9	4735	2 T17463	rifamycin polyketi
20	197	48.4	4613	2 T17409	polyketide synthas
21	196	48.2	1198	2 T28678	polyketide synthas
22	192.5	47.3	2723	2 T03221	probable polyketid
23	186	45.7	3519	2 S43078	polyketide synthas
24	180	44.2	3573	2 S23070	erythronolide synt
25	176	43.2	1346	2 T17412	polyketide synthas
26	172.5	42.4	1616	2 G70668	polyketide synthas
27	172	42.3	502	2 A70965	probable polyketid
28	169	41.5	1402	2 D70634	probable polyketid
29	169	41.5	2103	2 G86925	probable polyketid

30	167.5	41.2	2126	2 H70621	probable polyketid
31	151	37.1	2124	2 T28658	polyketide synthas
32	139	34.2	2201	2 S73014	polyketide synthas
33	135.5	33.3	1017	2 B70985	probable polyketid
34	122	30.0	2168	2 A70984	probable polyketid
35	117	28.7	8243	2 T31307	type I fatty acid
36	113	27.8	2478	2 AH2140	polyketide synthas
37	104.5	25.7	4427	2 PN0637	polyketide synthet
38	103.5	25.4	4930	2 E69679	hypothetical prote
39	101.5	24.9	7829	2 T15789	probable polyketid
40	99.5	24.4	2458	2 T17420	polyketide synthas
41	99	24.3	1540	2 H87203	probable ppsb prot
42	98	24.1	1538	2 E70874	polyketide synthas
43	97.5	24.0	958	2 S73012	hypothetical prote
44	97.5	24.0	1570	2 AC2012	polyketide synthas
45	97.5	24.0	1871	2 A87204	

ALIGNMENTS

RESULT 1
T17410
polyketide synthase type I - Streptomyces venezuelae
C/Species: Streptomyces venezuelae
C/Date: 02-Sep-2000 #sequence_reviseion 02-Sep-2000 #text_change 17-Nov-2000
C/Accession: T17410
R/Xue, Y., Zhao, L., Liu, H.W., Sherman, D.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
A/Title: A gene cluster for macrolide antibiotic biosynthesis in streptomyces venezuela
A/Reference number: Z18773; PMID:98445333; PMID:9770448
A/Accession: T17410
A/Status: preliminary; translated from GB/EMBL/DBD3
A/Molecule type: DNA
A/Residues: 1-3739 <XUE>
A/Cross-references: EMBL:AF079138; NID:G3808326; PID:G3808335; PIDN:AAC69330.1
C/Genetics:
A/Genes: pikAII
C/Keywors: antibiotic biosynthesis; carrier protein
F/1445-1516/Domain: acyl carrier protein homology <ACPI>
F/3570-3641/Domain: acyl carrier protein homology <ACP2>

Query Match 56.3%; Score 229; DB 2; Length 3739;
Best Local Similarity 59.8%; Pred. No. 8.9e-17;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;

Db 1 LAHQTSADQRAALLEVRDVAALRHADPKAIPDQSPFALGFDLSITAVEFRNLTKAT 60
3560 LAAMTPDDRVAVHLRDLVRGTVATVGLHGTPSRVLDLRAPRDTGFDLSITAVELNNRLNAAAT 3619
QY 61 GRLPVSIVFDHPTPAKLAVHL 82
Db 3620 GRLPVSIVFDHPTPAKLAVHL 3641

RESULT 2
T17411
polyketide synthase III - Streptomyces venezuelae
C/Species: Streptomyces venezuelae
C/Date: 02-Sep-2000 #sequence_reviseion 02-Sep-2000 #text_change 03-Nov-2000
C/Accession: T17411
R/Xue, Y., Zhao, L., Liu, H.W., Sherman, D.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
A/Title: A gene cluster for macrolide antibiotic biosynthesis in streptomyces venezuela
A/Reference number: Z18773; PMID:98445333; PMID:9770448
A/Accession: T17411
A/Status: preliminary; translated from GB/EMBL/DBD3
A/Molecule type: DNA
A/Residues: 1-1562 <XUE>
A/Cross-references: EMBL:AF079138; NID:G3808326; PID:G3808335; PIDN:AAC69331.1
C/Genetics:
A/Genes: pikAIII

F:5307-5378/Domain: acyl carrier protein homology <ACP2>
 F:5431-5831/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
 F:5532-6206/Domain: acyl carrier protein S-malonyltransferase homology <AMT2>
 F:6293-6364/Domain: acyl carrier protein homology <ACP3>

Query Match 54.3%; Score 221; DB 2; Length 6420;
 Best Local Similarity 57.3%; Pred. No. 1.3e-15;
 Matches 47; Conservative 9; Mismatches 26; Indels 0; Gaps 0;
 Oy 1 LAHQTSDQRAALLLEVRDHVAALRHADPKAIAPDQSPFALGFDLSITAVEFRNLTKAT 60
 Db 3381 LAELPADRDPTALTVDSTAAVAGHADASGIAPTTFKDLGIDSLTALELRNLLEAT 3440

Gy 61 GLRLPVSIVDFDHPPTPAKLAHVL 82
 Db 3441 GLRLSATMVFDPHPTPRVLAHVL 3462

RESULT 7
 T03222
 Probable polyketide synthase module 2 - Streptomyces hygroscopicus
 C:Species: Streptomyces hygroscopicus
 C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 03-Nov-2000
 C/Accession: T03222
 R:Ruan, X.; Staehli, D.; Lax, S.; Katz, L.
 Gene 203, 1-9, 1997
 A>Title: A second type-I PKS gene cluster isolated from Streptomyces hygroscopicus ATCC
 A/Reference number: Z14848; MUID:98085969; PMID:9426000
 A/Accession: T03222
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1762 <RNA>
 A/Cross-references: EMBL:AF007101; NID:G62424946; PIDN:AA38062.1; PID:G2624949
 A/Experimental source: ATCC 29253
 C:Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoac
 C/Keywords: [acyl-carrier-protein] S-malonyltransferase homology
 C/Keywords: carrier protein
 F:54-54/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
 F:550-822/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
 F:1598-1669/Domain: acyl carrier protein homology <ACP>

Query Match 53.2%; Score 216.5; DB 2; Length 1762;
 Best Local Similarity 56.5%; Pred. No. 9.6e-16;
 Matches 48; Conservative 8; Mismatches 20; Indels 9; Gaps 1;

Oy 7 ADQRAAL-----LELYRDHVAALVRHADPKAIAPDQSPFALGFDLSITAVEFRNL 57
 Db 1585 ADRIALMPERDRAVLDVROQSAFVLSGSDADRIDPQAFKDTGFDLSITAVEFRNRT 1644
 Gy 58 KATGLRLPVSIVDFDHPPTPAKLAHVL 82
 Db 1645 TATGLRLPATLIFDHPPTNALADHL 1669

RESULT 8
 T17467
 rifamycin polyketide synthase modules 9-10 - Amycolatopsis mediterranei
 C:Species: Amycolatopsis mediterranei
 C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
 C/Accession: T17467
 R:Schupp, T.
 A/Submitted to the EMBL Data Library, December 1997
 A/Reference number: Z18802
 A/Accession: T17467
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-3413 <SCH>
 A/Cross-references: EMBL:AF223012; NID:e1227119; PID:e1227124; PIDN:CAA11039.1
 A/Experimental source: strain LBG A336
 C:Superfamily: acyl carrier protein homology
 C/Keywords: carrier protein
 F:1608-1679/Domain: acyl carrier protein homology <ACP1>
 F:3334-3405/Domain: acyl carrier protein homology <ACP2>

Query Match 53.1%; Score 216; DB 2; Length 3413;
 Best Local Similarity 54.9%; Pred. No. 2.3e-15;
 Matches 45; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

Oy 1 LAHQTSDQRAALLLEVRDHVAALRHADPKAIAPDQSPFALGFDLSITAVEFRNLTKAT 60
 Db 1598 LAELPADRDPTALTVDSTAAVAGHADASGIAPTTFKDLGIDSLTALELRNLLEAT 3440
 Gy 61 GLRLPVSIVDFDHPPTPAKLAHVL 82
 Db 1658 GLRLPATLIVDFDHPPTPAKLAHVL 1679

RESULT 9
 G70944
 Probable polyketide synthase Rv2048c - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C/Accession: G70944
 R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; MUID:9825987; PMID:9634230
 A/Accession: G70944
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-4151 <COL>
 A/Cross-references: GB:AL021899; GB:AL123456; NID:G3242282; PIDN:CAA17262.1; PID:G28967
 A/Experimental source: strain H37RV
 C:Genetics:
 A:Gene: pxe12
 C:Superfamily: Mycobacterium tuberculosis probable polyketide synthase Rv2048c; 3-oxoac
 C/Keywords: [acyl-carrier-protein] S-malonyltransferase homology
 C/Keywords: carrier protein
 F:56-455/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
 F:559-841/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
 F:1680-1861/Domain: short-chain alcohol dehydrogenase homology <SAD1>
 F:1963-2038/Domain: acyl carrier protein homology <ACP1>
 F:2078-2478/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
 F:2582-2860/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
 F:3710-3891/Domain: short-chain alcohol dehydrogenase homology <SAD2>
 F:3895-4066/Domain: acyl carrier protein homology <ACP2>

Query Match 53.1%; Score 216; DB 2; Length 4151;
 Best Local Similarity 54.7%; Pred. No. 2.8e-15;
 Matches 47; Conservative 14; Mismatches 21; Indels 4; Gaps 1;

Oy 1 LAHQ---TSADQRAALLLEVRDHVAALRHADPKAIAPDQSPFALGFDLSITAVEFRNL 56
 Db 1949 LAHRLHGLPEABEQHVAALVGLVRLHATVAGNITPEAIDPDKAFQDGLFDSITAVEFRNL 2008
 Gy 57 IKATGLRLPVSIVDFDHPPTPAKLAHVL 82
 Db 2009 KSAIGLSLSPTLIFDHPPTNRLASYI 2034

RESULT 10
 T30226
 polyketide synthase - Streptomyces hygroscopicus
 C:Species: Streptomyces hygroscopicus
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 01-Dec-2000
 C/Accession: T30226
 R:Aparicio, J.F.; Molinar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Kraw, L.E.; Stau
 Gene 169, 9-16, 1996
 A>Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hy
 A/Reference number: Z20782; MUID:96186886; PMID:8635756
 A/Accession: T30226
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA

A;Residues: 1-8563 <APA>
A;Cross-references: EMBL:X86780; NID:9987088; PID:9987100; PIDN:CAA60460.1
C;Genetics:
A;Gene: rapA
C;Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase
C;Keywords: carrier protein; phosphopantetheine; phosphoprotein
F;54-503/Domain: acetate-CoA ligase homology <ACLI>
F;1329-1724/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F;1817-2091/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F;3199-3270/Domain: acyl carrier protein homology <ACP1>
F;3314-3706/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F;4787-4858/Domain: acyl carrier protein homology <ACP2>
F;4902-5293/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F;5386-5659/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F;6760-6831/Domain: acyl carrier protein homology <ACP3>
F;6875-7269/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F;7362-7638/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
F;8412-8483/Domain: acyl carrier protein homology <ACP4>
F;4822/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 52.8%; Score 215; DB 2; Length 8563;
Best Local Similarity 58.5%; Pred. No. 8.2e-15;
Matches 46; Conservative 7; Mismatches 27; Indels 0; Gaps 0;

QY 1 LAHQTADQRAALDELVRDVAALVRADPKAIPDQSPFALGDSITAVEFNLLIKAT 60
Db 4777 LAALAPAREKALKIKVCDASATVLGHADSTSVAAVFNRLGVDLSITAVEFNLSLAKAT 4836

QY 61 GLRLPVSIVFDHPPTPAKLAHVL 82
Db 4837 GLRLPATVFDYPTPTALAVRL 4858

RESULT 11

T30225
polyketide synthase - Streptomyces hygroscopicus
C;Species: Streptomyces hygroscopicus
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
C;Accession: T30225
R;Apicario, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Staun
Gene 169, 9-16, 1996
A;Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hyg
A;Reference number: Z20782; MUID:96186896; PMID:8635756
A;Accession: T30225
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-10223 <APA>
A;Cross-references: EMBL:X86780; NID:9987088; PID:9987099; PIDN:CAA60459.1
C;Genetics:
A;Gene: rapB
C;Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protei
C;Keywords: carrier protein
F;54-449/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F;1533-1604/Domain: acyl carrier protein homology <ACP1>
F;1647-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F;2135-2409/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F;3172-3243/Domain: acyl carrier protein homology <ACP2>
F;3287-3681/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F;3778-4052/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F;5160-5231/Domain: acyl carrier protein homology <ACP3>
F;5375-5661/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F;6776-6841/Domain: acyl carrier protein homology <ACP4>
F;6691-7285/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS5>
F;8411-8482/Domain: acyl carrier protein homology <ACP5>
F;8526-8921/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS6>
F;9012-9285/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
F;10069-10140/Domain: acyl carrier protein homology <ACP>

Query Match 52.7%; Score 214.5; DB 2; Length 10223;
Best Local Similarity 60.5%; Pred. No. 1.1e-14;
Matches 46; Conservative 11; Mismatches 18; Indels 1; Gaps 1;

QY 8 DQRA-ALLELVDRDVAALVRADPKAIPDQSPFALGDSITAVEFNLLIKATGLRLP 66

Db 10065 EERAKALKVRRDSATVLGHADRSIPATGAFKGLGVDLSITAVEFNLSLTKATGLRLPA 10124
QY 67 SLVFDHPPTPAKLAHVL 82
Db 10125 TWFVDFPTPADLAARL 10140

RESULT 12

Tifamycin polyketide synthase module 8 - Amycolatopsis mediterranei
C;Species: Amycolatopsis mediterranei
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 03-Nov-2000
C;Accession: T17466
R;Schupp, T.
submitted to the EMBL Data Library, December 1997

A;Reference number: Z18602
A;Accession: T17466
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1728 <SCH>

A;Cross-references: EMBL:A222012; NID:e1227119; PID:e1227123; PIDN:CAA11038.1
A;Experimental source: strain DBG A3136
C;Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoac
C;Keywords: carrier protein
F;1573-1644/Domain: acyl carrier protein homology <ACP>

Query Match 52.1%; Score 212; DB 2; Length 1728;
Best Local Similarity 53.7%; Pred. No. 3e-15;
Matches 44; Conservative 14; Mismatches 24; Indels 0; Gaps 0;

QY 1 LAHQTADQRAALDELVRDVAALVRADPKAIPDQSPFALGDSITAVEFNLLIKAT 60
Db 1563 LAGCGAPEQDALLVDLRGVAAVLRGADPAVRAPTAFKADGFSITSDLRNRLREST 1622

QY 61 GLRLPVSIVFDHPPTPAKLAHVL 82
Db 1623 GLRLPATVLDYPTPTALAVRL 1644

RESULT 13

T30228
polyketide synthase - Streptomyces hygroscopicus
C;Species: Streptomyces hygroscopicus
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
C;Accession: T30228
R;Apicario, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Staun
Gene 169, 9-16, 1996
A;Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hy
A;Reference number: Z20782; MUID:96186896; PMID:8635756
A;Accession: T30228
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-6260 <APA>
A;Cross-references: EMBL:X86780; NID:9987088; PID:9987102; PIDN:CAA60462.1
C;Genetics:
A;Gene: rapC
C;Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier prote
C;Keywords: carrier protein
F;55-451/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F;1536-1607/Domain: acyl carrier protein homology <ACP1>
F;1647-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F;3123-3194/Domain: acyl carrier protein homology <ACP2>
F;3328-3633/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F;3735-3999/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F;5114-5185/Domain: acyl carrier protein homology <ACP3>
F;5242-5638/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F;5759-6044/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F;6135-6206/Domain: acyl carrier protein homology <ACP4>

Query Match 52.1%; Score 212; DB 2; Length 6260;
Best Local Similarity 58.5%; Pred. No. 1.3e-14;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 13:08:02 ; Search time 2.55496 Seconds
(without alignments)
3431.399 Million cell updates/sec

Title: US-09-914-286-3_COPY_1979_2060

Perfect score: 407
Sequence: 1 LAHQTSADQRAALLETVRDH.....RLPVSLVDPHTPAKLAHL 82

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications_AA:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCRUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	229	56.3	3739	9 US-09-860-846-33	Sequence 33, Appl
2	229	56.3	3739	9 US-09-988-384B-33	Sequence 33, Appl
3	229	56.3	3739	9 US-09-836-821-33	Sequence 33, Appl
4	229	56.3	3739	9 US-09-793-708-2	Sequence 2, Appl
5	229	56.3	3739	10 US-09-861-289-33	Sequence 33, Appl
6	229	56.3	11877	9 US-09-860-846-6	Sequence 6, Appl
7	229	56.3	11877	9 US-09-836-821-6	Sequence 6, Appl
8	229	56.3	11877	10 US-09-861-289-6	Sequence 6, Appl
9	229	56.3	12199	9 US-09-988-384B-6	Sequence 6, Appl
10	229	55.3	1562	9 US-09-860-846-35	Sequence 35, Appl
11	229	55.3	1562	9 US-09-988-384B-35	Sequence 35, Appl
12	229	55.3	1562	9 US-09-836-821-35	Sequence 35, Appl
13	229	55.3	1562	10 US-09-793-708-3	Sequence 3, Appl
14	229	55.3	1562	10 US-09-861-289-35	Sequence 35, Appl
15	229	55.0	3816	9 US-09-808-880-3	Sequence 3, Appl
16	229	54.8	4150	9 US-09-808-880-2	Sequence 2, Appl
17	229	51.8	5215	9 US-09-860-846-2	Sequence 2, Appl
18	229	51.8	5215	9 US-09-988-384B-2	Sequence 2, Appl
19	229	51.8	5215	9 US-09-836-821-2	Sequence 2, Appl

20	211	51.8	5215	10 US-09-861-289-2	Sequence 2, Appl
21	197	48.4	4551	9 US-09-793-708-1	Sequence 1, Appl
22	197	48.4	4613	9 US-09-860-846-31	Sequence 31, Appl
23	197	48.4	4613	9 US-09-988-384B-31	Sequence 31, Appl
24	197	48.4	4613	9 US-09-836-821-31	Sequence 31, Appl
25	197	48.4	4613	10 US-09-861-289-31	Sequence 31, Appl
26	186	45.7	3519	9 US-09-808-880-4	Sequence 4, Appl
27	178	43.7	1346	9 US-09-793-708-4	Sequence 4, Appl
28	176	43.2	1346	9 US-09-860-846-37	Sequence 37, Appl
29	176	43.2	1346	9 US-09-988-384B-37	Sequence 37, Appl
30	176	43.2	1346	9 US-09-836-821-37	Sequence 37, Appl
31	176	43.2	1346	10 US-09-861-289-37	Sequence 37, Appl
32	172.5	42.4	1616	9 US-09-712-363-262	Sequence 262, App
33	169	41.5	1402	9 US-09-712-363-166	Sequence 166, App
34	163	40.0	3798	9 US-10-014-717-6	Sequence 6, Appl
35	163	40.0	7257	9 US-10-014-717-5	Sequence 5, Appl
36	112	27.5	1832	9 US-10-014-717-4	Sequence 4, Appl
37	101	24.8	352	10 US-09-073-009-14	Sequence 14, Appl
38	101	24.8	352	10 US-09-023-588-14	Sequence 14, Appl
39	101	24.8	352	10 US-09-793-306-14	Sequence 14, Appl
40	98.5	24.2	2439	9 US-10-014-717-7	Sequence 7, Appl
41	97.5	24.0	1421	9 US-10-014-717-2	Sequence 2, Appl
42	86	21.1	1827	9 US-09-712-363-261	Sequence 261, App
43	85.5	21.0	1812	10 US-09-775-938A-38	Sequence 38, Appl
44	78.5	19.3	4999	9 US-09-976-059-15	Sequence 15, Appl
45	77	18.9	1051	9 US-09-976-059-13	Sequence 13, Appl

ALIGNMENTS

```

RESULT 1
US-09-860-846-33
; Sequence 33, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Xue, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438051
; CURRENT APPLICATION NUMBER: US/09/860,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 3739
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-33

Query Match          56.3%; Score 229; DB 9; Length 3739;
Best Local Similarity 59.8%; Pred. No. 1,1e-18;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;

QY      1 LAHQTSADQRAALLETVRDVAALVRHADKALAPQSPFALGPDSTLTAVERNLTKAT 60
DB      3560 LAAMTPDPRVAHLRDIVRTTVAIVLGHGTSRVDLBAFADTGFDSITAVELNRLMAAT 3619
QY      61 GLRLPVSLVDPHTPAKLAHL 82
DB      3620 GLRLPVSLVDPHTPAKLAHL 3641

RESULT 2
US-09-988-384B-33
; Sequence 33, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.

```

```
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.536US1
CURRENT APPLICATION NUMBER: US/09/988.384B
CURRENT FILING DATE: 2001-11-19
PRIORITY FILING DATE: PCT/US99/14398
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 53
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-988-384B-33
```

```
Query Match 56.3%; Score 229; DB 9; Length 3739;
Best Local Similarity 59.8%; Pred. No. 1.1e-18;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;
```

```
QY 1 LAHOTSADQRAALLEYVDHVAALVRHADPKAIPADPSFRALGDSLTAVEFNLLIKAT 60
DB 3560 LAAMTPDDRVAHLRDLVTRHVAATVLGHTPSRVDLERAFRDTGDSLTAVEFNRLNAAT 3619
GLRLPVSIVFDHPTPAKLAVHL 82
3620 GLRLPATIVFDHPTPGELAGHL 3641
```

```
RESULT 3
US-09-836-821-33
Sequence 33, Application US/09836821
Publication NO. US20030087405A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/836.821
CURRENT FILING DATE: 2001-04-17
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 09/105,537
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-836-821-33
```

```
Query Match 56.3%; Score 229; DB 9; Length 3739;
Best Local Similarity 59.8%; Pred. No. 1.1e-18;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;
```

```
QY 1 LAHOTSADQRAALLEYVDHVAALVRHADPKAIPADPSFRALGDSLTAVEFNLLIKAT 60
DB 3560 LAAMTPDDRVAHLRDLVTRHVAATVLGHTPSRVDLERAFRDTGDSLTAVEFNRLNAAT 3619
GLRLPVSIVFDHPTPAKLAVHL 82
3620 GLRLPATIVFDHPTPGELAGHL 3641
```

```
RESULT 4
US-09-793-708-2
Sequence 2, Application US/09793708
Publication No. US20030104597A1
GENERAL INFORMATION:
APPLICANT: Ashley, Gary
```

```
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002121
CURRENT APPLICATION NUMBER: US/09/793.708
CURRENT FILING DATE: 2001-02-22
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: US 09/320,878
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: US 09/141,908
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: US 09/073,538
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: US 08/846,247
PRIOR FILING DATE: 1997-04-30
PRIOR APPLICATION NUMBER: US 60/134,990
PRIOR FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-793-708-2
```

```
Query Match 56.3%; Score 229; DB 9; Length 3739;
Best Local Similarity 59.8%; Pred. No. 1.1e-18;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;
```

```
QY 1 LAHOTSADQRAALLEYVDHVAALVRHADPKAIPADPSFRALGDSLTAVEFNLLIKAT 60
DB 3560 LAAMTPDDRVAHLRDLVTRHVAATVLGHTPSRVDLERAFRDTGDSLTAVEFNRLNAAT 3619
GLRLPVSIVFDHPTPAKLAVHL 82
3620 GLRLPATIVFDHPTPGELAGHL 3641
```

```
RESULT 5
US-09-861-289-33
Sequence 33, Application US/09861289
Patent No. US20020110897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/861.289
CURRENT FILING DATE: 2001-05-18
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 09/105,537
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-861-289-33
```

```
Query Match 56.3%; Score 229; DB 10; Length 3739;
Best Local Similarity 59.8%; Pred. No. 1.1e-18;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;
```

```
QY 1 LAHOTSADQRAALLEYVDHVAALVRHADPKAIPADPSFRALGDSLTAVEFNLLIKAT 60
DB 3560 LAAMTPDDRVAHLRDLVTRHVAATVLGHTPSRVDLERAFRDTGDSLTAVEFNRLNAAT 3619
GLRLPVSIVFDHPTPAKLAVHL 82
```


Db 3620 GLRLPATLVFDHPTPGELAGHL 3641

RESULT 6
US-09-860-846-6

Sequence 6, Application US/09860846
Patent No. US2002016742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/860,846
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-860-846-6

Query Match 56.3%; Score 229; DB 9; Length 11877;
Best Local Similarity 59.8%; Pred. No. 5.1e-18;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;

Qy 1 LAHQTSAOQRAALLETVRDHVAALVRHADPKAIAPDQSPFALGFDSTLTAVERNLITKAT 60
Db 8509 LAAMTPDDRVAHLRDVLRTHVATVLCGHTPSRVDLERAFPDGTGDSLTAVELNRRLNAAT 8568
Qy 61 GLRLPVSLVFDHPTPAKLA VHL 82
Db 8569 GLRLPATLVFDHPTPGELAGHL 8590

RESULT 7

US-09-836-821-6
Sequence 6, Application US/09836821
Publication No. US20030087405A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/836,821
CURRENT FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-836-821-6

Query Match 56.3%; Score 229; DB 9; Length 11877;
Best Local Similarity 59.8%; Pred. No. 5.1e-18;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;

Qy 1 LAHQTSAOQRAALLETVRDHVAALVRHADPKAIAPDQSPFALGFDSTLTAVERNLITKAT 60
Db 8509 LAAMTPDDRVAHLRDVLRTHVATVLCGHTPSRVDLERAFPDGTGDSLTAVELNRRLNAAT 8568
Qy 61 GLRLPVSLVFDHPTPAKLA VHL 82

Db 8569 GLRLPATLVFDHPTPGELAGHL 8590

RESULT 8
US-09-861-289-6

Sequence 6, Application US/09861289
Patent No. US20020110897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/861,289
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 11877
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-861-289-6

Query Match 56.3%; Score 229; DB 10; Length 11877;
Best Local Similarity 59.8%; Pred. No. 5.1e-18;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;

Qy 1 LAHQTSAOQRAALLETVRDHVAALVRHADPKAIAPDQSPFALGFDSTLTAVERNLITKAT 60
Db 8509 LAAMTPDDRVAHLRDVLRTHVATVLCGHTPSRVDLERAFPDGTGDSLTAVELNRRLNAAT 8568
Qy 61 GLRLPVSLVFDHPTPAKLA VHL 82
Db 8569 GLRLPATLVFDHPTPGELAGHL 8590

RESULT 9

US-09-988-384B-6
Sequence 6, Application US/09988384B
Publication No. US20030073824A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.536US1
CURRENT APPLICATION NUMBER: US/09/988,384B
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: PCT/US99/14398
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 53
SEQ ID NO 6
LENGTH: 12199
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-988-384B-6

Query Match 56.3%; Score 229; DB 9; Length 12199;
Best Local Similarity 59.8%; Pred. No. 5.3e-18;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;

Qy 1 LAHQTSAOQRAALLETVRDHVAALVRHADPKAIAPDQSPFALGFDSTLTAVERNLITKAT 60
Db 8831 LAAMTPDDRVAHLRDVLRTHVATVLCGHTPSRVDLERAFPDGTGDSLTAVELNRRLNAAT 8890
Qy 61 GLRLPVSLVFDHPTPAKLA VHL 82

Db 8891 GLRLPALTIVFDHPPTGELAGHL 8912

RESULT 10
US-09-860-846-35
; Sequence 35, Application US/09860846
; Patent No. US2002016472A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 1562
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-35

Query Match 55.3%; Score 225; DB 9; Length 1562;
Best Local Similarity 60.0%; Pred. No. 1,1e-18;
Matches 45; Conservative 8; Mismatches 22; Indels 0; Gaps 0;

QY 8 DORAALELVVDHVAALVRHADPKAIAPOSFRLGFSITLAVFERNLLIKATGLRLPVS 67
DB 1400 ERRPALTLVTRTHAAAVLGHSSPDRVAPGAFTELGFSITLAVQLNQLSTVGNRLPAT 1459

QY 68 LVFDHPTPAKLAHVL 82
DB 1460 TVFDHPTPAALAHVL 1474

RESULT 11
US-09-988-384B-35
; Sequence 35, Application US/09988384B
; Publication No. US2003007824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 35
; LENGTH: 1562
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-35

Query Match 55.3%; Score 225; DB 9; Length 1562;
Best Local Similarity 60.0%; Pred. No. 1,1e-18;
Matches 45; Conservative 8; Mismatches 22; Indels 0; Gaps 0;

QY 8 DORAALELVVDHVAALVRHADPKAIAPOSFRLGFSITLAVFERNLLIKATGLRLPVS 67
DB 1400 ERRPALTLVTRTHAAAVLGHSSPDRVAPGAFTELGFSITLAVQLNQLSTVGNRLPAT 1459

QY 68 LVFDHPTPAKLAHVL 82

Db 1460 TVFDHPTPAALAHVL 1474

RESULT 12
US-09-836-821-35
; Sequence 35, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 1562
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-836-821-35

Query Match 55.3%; Score 225; DB 9; Length 1562;
Best Local Similarity 60.0%; Pred. No. 1,1e-18;
Matches 45; Conservative 8; Mismatches 22; Indels 0; Gaps 0;

QY 8 DORAALELVVDHVAALVRHADPKAIAPOSFRLGFSITLAVFERNLLIKATGLRLPVS 67
DB 1400 ERRPALTLVTRTHAAAVLGHSSPDRVAPGAFTELGFSITLAVQLNQLSTVGNRLPAT 1459

QY 68 LVFDHPTPAKLAHVL 82
DB 1460 TVFDHPTPAALAHVL 1474

RESULT 13
US-09-793-708-3
; Sequence 3, Application US/09793708
; Publication No. US20030104597A1
; GENERAL INFORMATION:
; APPLICANT: Ashley, Gary
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002121
; CURRENT APPLICATION NUMBER: US/09/793,708
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 09/657,440
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: US 09/141,908
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: US 09/073,538
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: US 08/846,247
; PRIOR FILING DATE: 1997-04-30
; PRIOR APPLICATION NUMBER: US 60/134,990
; PRIOR FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1562
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-793-708-3

Query Match 55.3%; Score 225; DB 9; Length 1562;
Best Local Similarity 60.0%; Pred. No. 1.1e-18;
Matches 45; Conservative 8; Mismatches 22; Indels 0; Gaps 0;

QY 8 DORALLLEVRDVAALVLRADPKAIPDGSFRLGFSLTAVEFRNLLIKATGLRLPVS 67
DB 1400 ERRALLTLVTRTHAAALVGHSSPDVAVPAGRAFTELGFSLTAVQLRNQLSTVVGNRLLPAT 1459

QY 68 LVFDHPTPAKLAAYHL 82
DB 1460 TVFDHPTPALAAHL 1474

RESULT 14

US-09-861-289-35
Sequence 35, Application US/09861289
Patent No. US20020110897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438US1
CURRENT APPLICATION NUMBER: US/09/861,289
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 35
LENGTH: 1562
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-861-289-35

Query Match 55.3%; Score 225; DB 10; Length 1562;
Best Local Similarity 60.0%; Pred. No. 1.1e-18;
Matches 45; Conservative 8; Mismatches 22; Indels 0; Gaps 0;

QY 8 DORALLLEVRDVAALVLRADPKAIPDGSFRLGFSLTAVEFRNLLIKATGLRLPVS 67
DB 1400 ERRALLTLVTRTHAAALVGHSSPDVAVPAGRAFTELGFSLTAVQLRNQLSTVVGNRLLPAT 1459

QY 68 LVFDHPTPAKLAAYHL 82
DB 1460 TVFDHPTPALAAHL 1474

RESULT 15

US-09-808-880-3
Sequence 3, Application US/09808880
Publication No. US20030027287A1
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Kishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/808,880
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: US/09/428,517
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/120,254
PRIOR FILING DATE: 1999-02-16
PRIOR APPLICATION NUMBER: 60/106,100
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 3816
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: Oleandolide PKS
US-09-808-880-3

Query Match 55.0%; Score 224; DB 9; Length 3816;
Best Local Similarity 62.7%; Pred. No. 4.7e-18;
Matches 47; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

QY 8 DORALLLEVRDVAALVLRADPKAIPDGSFRLGFSLTAVEFRNLLIKATGLRLPVS 67
DB 3632 DRRRALRDVLRGTHALVLRGSHGSEPAIDARQAFRDIGFSLTAVELRNRLNAFTGLRLPGT 3691

QY 68 LVFDHPTPAKLAAYHL 82
DB 3692 LVFDYRNPALADHL 3706

Search completed: June 17, 2003, 13:46:43
Job time : 2.55496 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:54:32 ; Search time 1.49039 Seconds
(without alignments)
1618.821 Million cell updates/sec

Title: US-09-914-286-3_COPY_1979_2060

Perfect score: 407

Sequence: 1 LAHQTSADQRAALIEVRDH.....RLPVSIVFDHPTPAKLAVHL 82

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	229	56.3	3739	3 US-09-320-878-2	Sequence 2, Appl
2	229	56.3	3739	4 US-09-105-537-3	Sequence 3, Appl
3	229	56.3	11877	4 US-09-105-537-6	Sequence 6, Appl
4	225	55.3	1562	3 US-09-320-878-3	Sequence 3, Appl
5	225	55.3	1562	4 US-09-105-537-35	Sequence 35, Appl
6	224	55.0	3816	4 US-09-428-517-3	Sequence 3, Appl
7	223	54.8	3724	2 US-08-804-227C-10	Sequence 10, Appl
8	223	54.8	3724	2 US-08-804-198-4	Sequence 4, Appl
9	223	54.8	4150	4 US-09-428-517-2	Sequence 2, Appl
10	222	54.5	4545	2 US-08-804-227C-14	Sequence 14, Appl
11	222	54.5	4550	2 US-08-804-227C-8	Sequence 8, Appl
12	222	54.5	4550	2 US-08-804-198-2	Sequence 2, Appl
13	219	53.8	1996	2 US-08-804-227C-9	Sequence 9, Appl
14	219	53.8	1996	2 US-08-804-198-3	Sequence 3, Appl
15	219	53.8	3729	2 US-08-804-227C-4	Sequence 4, Appl
16	212	52.1	2152	4 US-09-036-987A-3	Sequence 3, Appl
17	212	52.1	2152	4 US-09-370-700-3	Sequence 3, Appl
18	211	51.8	5215	4 US-09-105-537-2	Sequence 2, Appl
19	210	51.6	3170	4 US-09-036-987A-4	Sequence 4, Appl
20	210	51.6	3170	4 US-09-370-700-4	Sequence 4, Appl
21	209.5	51.5	185	4 US-09-154-083-14	Sequence 14, Appl
22	205	50.4	3170	2 US-07-642-734C-5	Sequence 5, Appl
23	205	50.4	3170	3 US-08-439-009A-5	Sequence 5, Appl
24	205	50.4	3567	2 US-07-642-734C-4	Sequence 4, Appl
25	205	50.4	3567	3 US-08-439-009A-4	Sequence 4, Appl
26	202	49.6	3491	2 US-07-642-734C-2	Sequence 2, Appl
27	202	49.6	3491	3 US-08-439-009A-2	Sequence 2, Appl

28	197	48.4	4551	3 US-09-320-878-1	Sequence 1, Appl
29	197	48.4	4613	4 US-09-105-537-31	Sequence 31, Appl
30	195	47.9	4928	4 US-09-036-987A-5	Sequence 5, Appl
31	195	47.9	4928	4 US-09-370-700-5	Sequence 5, Appl
32	191	46.9	5588	4 US-09-036-987A-6	Sequence 6, Appl
33	191	46.9	5588	4 US-09-370-700-6	Sequence 6, Appl
34	186	45.7	1864	2 US-08-804-227C-3	Sequence 3, Appl
35	186	45.7	3519	4 US-09-428-517-4	Sequence 4, Appl
36	183	45.0	4472	2 US-08-804-227C-2	Sequence 2, Appl
37	182	44.7	502	4 US-09-413-814-106	Sequence 106, App
38	178	43.7	1346	3 US-09-320-878-4	Sequence 4, Appl
39	178	43.7	2595	4 US-09-036-987A-2	Sequence 2, Appl
40	178	43.7	2595	4 US-09-370-700-2	Sequence 2, Appl
41	176	43.2	1346	4 US-09-105-537-37	Sequence 37, Appl
42	176	43.2	1611	2 US-08-804-227C-5	Sequence 5, Appl
43	174	42.8	1841	2 US-08-804-227C-6	Sequence 6, Appl
44	165	40.5	1891	2 US-08-804-227C-12	Sequence 12, Appl
45	165	40.5	1891	2 US-08-804-198-6	Sequence 6, Appl

ALIGNMENTS

```

RESULT 1
US-09-320-878-2
; Sequence 2, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320, 878A
; EARLIER FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141, 908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073, 538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846, 247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119, 139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100, 880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087, 080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 3739
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-320-878-2

Query Match      56.3%; Score 229; DB 3; Length 3739;
Best Local Similarity 59.8%; Pred. No. 9, 8e-22;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;

QY      1 LAHQTSADQRAALIEVRDVAALVRADPKALAPDOSPALGDSITAVFNNLIKAT 60
Db      3560 LAAMTDDRAAHARLDVTRHVAIVLGHGTRSVDLERAFADDTGFSITAVELNRRLNAAT 3619

QY      61 GLRLPVSIVFDHPTPAKLAVHL 82
Db      3620 GLRLPATVIVFDHPTPELAGHL 3641

RESULT 2
US-09-105-537-33
; Sequence 33, Application US/09105537A

```

```

; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105.537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 3739
; TYPE: PR1
; ORGANISM: Streptomyces venezuelae
; US-09-105-537-33

Query Match          56.3%; Score 229; DB 4; Length 3739;
Best Local Similarity 59.8%; Pred. No. 9.8e-22;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;

QY 1 LAHGTSDQRAALIELVRDVAALRHADPKAIAPDOSFRALGDSLTAVEFRNLIKAT 60
DB 3560 LAAMTPDDRVAAHLKDLVTRHVAIVLGHGTPSRVLDLERAFTDGTGDSLTAVELRNRLNAAT 3619
QY 61 GLRLPVSIVPDHPPTPAKLAVAL 82
DB 3620 GLRLPATLVDPHPPTGELAGHL 3641

RESULT 3
US-09-105-537-6
; Sequence 6, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105.537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 11877
; TYPE: PR1
; ORGANISM: Streptomyces venezuelae
; US-09-105-537-6

Query Match          56.3%; Score 229; DB 4; Length 11877;
Best Local Similarity 59.8%; Pred. No. 5e-21;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;

QY 1 LAHGTSDQRAALIELVRDVAALRHADPKAIAPDOSFRALGDSLTAVEFRNLIKAT 60
DB 8509 LAAMTPDDRVAAHLKDLVTRHVAIVLGHGTPSRVLDLERAFTDGTGDSLTAVELRNRLNAAT 8568
QY 61 GLRLPVSIVPDHPPTPAKLAVAL 82
DB 8569 GLRLPATLVDPHPPTGELAGHL 8590

RESULT 4
US-09-320-878-3
; Sequence 3, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: Ashley, Gary
; APPLICANT: Betlach, Melanie C.
; APPLICANT: Betlach, Mary C.
```

```

; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320.878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1562
; TYPE: PR1
; ORGANISM: Streptomyces venezuelae
; US-09-320-878-3

Query Match          55.3%; Score 225; DB 3; Length 1562;
Best Local Similarity 60.0%; Pred. No. 1e-21;
Matches 45; Conservative 8; Mismatches 22; Indels 0; Gaps 0;

QY 8 DQRAALIELVRDVAALRHADPKAIAPDOSFRALGDSLTAVEFRNLIKATGLRLPVS 67
DB 1400 ERRPALTLTVTRHAAVNLGHSSPDVAPGSAFTLGTGDSLTAVOLRNQSLTVVGNRLPAT 1459
QY 68 LVFDHPPTPAKLAVAL 82
DB 1460 TVFDHPPTPAALAAHL 1474

RESULT 5
US-09-105-537-35
; Sequence 35, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105.537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 1562
; TYPE: PR1
; ORGANISM: Streptomyces venezuelae
; US-09-105-537-35

Query Match          55.3%; Score 225; DB 4; Length 1562;
Best Local Similarity 60.0%; Pred. No. 1e-21;
Matches 45; Conservative 8; Mismatches 22; Indels 0; Gaps 0;

QY 8 DQRAALIELVRDVAALRHADPKAIAPDOSFRALGDSLTAVEFRNLIKATGLRLPVS 67
DB 1400 ERRPALTLTVTRHAAVNLGHSSPDVAPGSAFTLGTGDSLTAVOLRNQSLTVVGNRLPAT 1459
QY 68 LVFDHPPTPAKLAVAL 82
DB 1460 TVFDHPPTPAALAAHL 1474
```

RESULT 6
US-09-428-517-3
Sequence 3, Application US/09428517
Patent No. 6251636
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/428,517
CURRENT FILING DATE: 1999-10-28
EARLIER APPLICATION NUMBER: 60/120,254
EARLIER FILING DATE: 1999-02-16
EARLIER APPLICATION NUMBER: 60/106,100
EARLIER FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 3
LENGTH: 3816
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-428-517-3
Query Match 55.0%; Score 224; DB 4; Length 3816;
Best Local Similarity 62.7%; Pred. No. 4,96-21;
Matches 47; Conservative 9; Mismatches 19; Indels 0; Gaps 0;
QY 8 DORALALEVRDHAALVLRHADPKAIAPOSPFALGFDSTLAVFRLIKATGLRPLVS 67
DB 3632 DPRALADLVGRHVALVGHSGPFAIDAEQAFRIGFDSLTAVALRRLNAETGLRLPGT 3691
QY 68 LVFDHPTPAKLAHL 82
DB 3692 LVFDYFNPSPALADHL 3706
RESULT 7
US-08-804-227C-10
Sequence 10, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhnle, Stuart A.
APPLICANT: Roestek, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 3724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-227C-10
Query Match 54.8%; Score 223; DB 2; Length 3724;
Best Local Similarity 54.9%; Pred. No. 6,4e-21;
Matches 45; Conservative 12; Mismatches 25; Indels 0; Gaps 0;
QY 1 LAHQTADQRAALLEVVDHVAALVRHADPKAIAPOSPFALGFDSTLAVFRLIKAT 60
DB 3548 LAGLSAEERHRLGLVGHEVAALVGHSAAEVRPDRPREVGFDSTLAVELRNMAAVT 3607
QY 61 GLRLPVSIVFDHPTPAKLAHL 82
DB 3608 GVRLPATLVFDHPTPALSSHL 3629
RESULT 8
US-08-804-198-4
Sequence 4, Application US/08804198
Patent No. 5945320
GENERAL INFORMATION:
APPLICANT: Burgett, Stanley G.
APPLICANT: Kuhnle, Stuart A.
APPLICANT: Rao, Nagaraia R.
APPLICANT: Richardson, Mark A.
APPLICANT: Roestek, Paul R., Jr.
TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-4
Query Match 54.8%; Score 223; DB 2; Length 3724;
Best Local Similarity 54.9%; Pred. No. 6,4e-21;
Matches 45; Conservative 12; Mismatches 25; Indels 0; Gaps 0;
QY 1 LAHQTADQRAALLEVVDHVAALVRHADPKAIAPOSPFALGFDSTLAVFRLIKAT 60
DB 3548 LAGLSAEERHRLGLVGHEVAALVGHSAAEVRPDRPREVGFDSTLAVELRNMAAVT 3607
QY 61 GLRLPVSIVFDHPTPAKLAHL 82

Db 3608 GVRLPATLVDPHPTPALSSHL 3629

RESULT 9

US-09-428-517-2
; Sequence 2, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Belach, Mary C.
; APPLICANT: Shan, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; EARLIER FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: Recombinant
; OTHER INFORMATION: Oleandolide PKS
US-09-428-517-2

Query Match 54.8%; Score 223; DB 4; Length 4150;
Best Local Similarity 56.1%; Pred. No. 7.5e-21;
Matches 46; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

QY 1 LAHOTSADORAALLELVRDHVAALRHADPKAIAPDOSFRALGFSLTAVEFRMLIKAT 60
DB 2441 LAGLTAVGQERLVLVQVREGAAVVLGHSAGAAVADPAFOLGFSLTAVEFRMLIKAT 2500

QY 61 GLRLPVLVDPHPTPAKLAHL 82
DB 2501 GLRLPVLVDPHPTPAKLAHL 2522

RESULT 10

US-08-804-227C-14
; Sequence 14, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 4545 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-804-227C-14

Query Match 54.5%; Score 222; DB 2; Length 4545;
Best Local Similarity 57.9%; Pred. No. 1.2e-20;
Matches 44; Conservative 10; Mismatches 22; Indels 0; Gaps 0;

QY 7 ADORAALELVRDHVAALRHADPKAIAPDOSFRALGFSLTAVEFRMLIKATGLRLPV 66
DB 2534 AERRRALTLVRESEAGILRHASADAVDPFLAFRSAGFDSLTVELRNRLTAATGLNLPN 2593

QY 67 SLVFDHPTPAKLAHL 82
DB 2594 TLVFDHPTPLSLASHL 2609

RESULT 11

US-08-804-227C-8
; Sequence 8, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 4550 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-227C-8

Query Match 54.5%; Score 222; DB 2; Length 4550;
Best Local Similarity 57.9%; Pred. No. 1.2e-20;
Matches 44; Conservative 10; Mismatches 22; Indels 0; Gaps 0;

QY 7 ADORAALELVRDHVAALRHADPKAIAPDOSFRALGFSLTAVEFRMLIKATGLRLPV 66

Db 2539 AERRRALLTVRSEBAGILRHASADAVDEPLAFRSAGFDSLTVLELNRLTAATGMLPN 2598
QY 67 SLVFDHPTPAKLAVHL 82
Db 2599 TLVFDHPTPLSLASHL 2614

RESULT 12
US-08-804-198-2
Sequence 2, Application US/08804198
Patent No. 5945320
GENERAL INFORMATION:
APPLICANT: Burgett, Stanley G.
APPLICANT: Kuhnloos, Stuart A.
APPLICANT: Rao, Nagareja R.
APPLICANT: Richardson, Mark A.
APPLICANT: Rosteck, Paul R., Jr.
TITLE OF INVENTION: PLATENOILIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PAUL R. CANTRELL, 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4550 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-2
Query Match 54.5%; Score 222; DB 2; Length 4550;
Best Local Similarity 57.9%; Pred. No. 1,2e-20;
Matches 44; Conservative 10; Mismatches 22; Indels 0; Gaps 0;
QY 7 ADOORALLLEVRDHVAALVRHADPKAIAPOSFRALGFDSLTVAFPNLLIKATGLPLV 66
Db 2539 AERRRALLTVRSEBAGILRHASADAVDEPLAFRSAGFDSLTVLELNRLTAATGMLPN 2598
QY 67 SLVFDHPTPAKLAVHL 82
Db 2599 TLVFDHPTPLSLASHL 2614
RESULT 13
US-08-804-227C-9
Sequence 9, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhnloos, Stuart A.
APPLICANT: Rosteck, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1996 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-227C-9
Query Match 53.8%; Score 219; DB 2; Length 1996;
Best Local Similarity 58.2%; Pred. No. 9.4e-21;
Matches 46; Conservative 11; Mismatches 22; Indels 0; Gaps 0;
QY 1 LAHQTSADQRAALLLEVRDHVAALVRHADPKAIAPOSFRALGFDSLTVAFPNLLIKAT 60
Db 1811 LAGRSAPERTALLLDVRTEVAALVGHGDPALGAARTFDMAGFDSLTVAVDLNRLNTRT 1870
QY 61 GLRLPVSLVFDHPTPAKLA 79
Db 1871 GLRLPATLVFDHPTPLALA 1889
RESULT 14
US-08-804-198-3
Sequence 3, Application US/08804198
Patent No. 5945320
GENERAL INFORMATION:
APPLICANT: Burgett, Stanley G.
APPLICANT: Kuhnloos, Stuart A.
APPLICANT: Rao, Nagareja R.
APPLICANT: Richardson, Mark A.
APPLICANT: Rosteck, Paul R., Jr.
TITLE OF INVENTION: PLATENOILIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PAUL R. CANTRELL, 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.

REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1996 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-3

Query Match 53.8%; Score 219; DB 2; Length 1996;
Best Local Similarity 58.2%; Pred. No. 9,4e-21;
Matches 46; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

OY 1 LAHQTADQRAALIELVDHVAALVLRHADPKAIAPDOSFRALGFDSTLTAVEFNNLIKAT 60
DB 1811 LAGRSAPERTALDLVTEVAALVGHGDPALIGARTFKDAGFDSLTAVDLRRRLNTRT 1870

OY 61 GLRLPVSIVFDHPTPAKLA 79
DB 1871 GLRLPATVFDHPTPALA 1889

RESULT 15

US-08-804-227C-4
Sequence 4, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: DeHoff, Bradley S.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rostreck, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3729 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-804-227C-4

Query Match 53.8%; Score 219; DB 2; Length 3729;
Best Local Similarity 57.7%; Pred. No. 2,3e-20;
Matches 45; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

OY 5 TSADQRAALIELVDHVAALVLRHADPKAIAPDOSFRALGFDSTLTAVEFNNLIKATGRL 64
DB 3551 TAOGHVAVLAEIRAAHAAVAVGHGSDSIDEDRAFKDLGFDSTLTAVEFNNLIKATGRL 3610

OY 65 PVSIVFDHPTPAKLAVAL 82
DB 3611 PATVFDHPTPGELAGHL 3628

Search completed: June 17, 2003, 13:09:28
Job time : 2.49039 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 12:50:57 ; Search time 4.01696 Seconds
(without alignments)
2720.102 Million cell updates/sec

Title: US-09-914-286-3_COPY_1979_2060

Perfect score: 407
Sequence: 1 LAHQTSADQRAALLLEVRDH.....RLPVSIVDFHPTAKLAVHL 82

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: A_GeneSeq_101002.*
2: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.*
3: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.*
4: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	407	100.0	3972	21	AA823749
2	407	100.0	3972	22	AA65264
3	407	100.0	3972	22	AA65268
4	243	59.7	5532	21	AA823752
5	243	59.7	5532	22	AA65267
6	240	59.0	6797	22	AA81558
7	233	57.2	6239	21	AA823750
8	233	57.2	6239	22	AA65265
9	229	56.3	3739	21	AA818638
10	229	56.3	3739	21	AA77193

11	229	56.3	3739	21	AA77201	S. venezuelae pik
12	229	56.3	3739	21	AA67202	Narboxonolide synthase
13	229	56.3	9477	22	AAE10144	Streptomycetes nous
14	229	56.3	12199	21	AA772180	S. venezuelae pik
15	225	55.3	1560	21	AA67203	Narboxonolide synthase
16	225	55.3	1561	21	AA77202	S. venezuelae pik
17	225	55.3	1562	21	AA818639	Amino acid sequenc
18	225	55.3	1562	21	AA77194	S. venezuelae macr
19	224	55.0	1721	19	AAW52847	A. mediterranei r1
20	224	55.0	3816	21	AA923708	S. antibioticus 8,
21	223	54.8	3724	18	AAW23718	Platenolide synthase
22	223	54.8	3724	18	AAW2608	Platenolide synthase
23	223	54.8	4150	21	AAW2707	S. antibioticus 8,
24	222	54.5	4545	18	AAW2611	Hybrid erm/tylg O
25	222	54.5	4550	18	AAW23716	Platenolide synthase
26	222	54.5	4550	18	AAW2606	Platenolide synthase
27	221	54.3	5069	19	AAW52846	A. mediterranei r1
28	221	54.3	11096	22	AAE10129	Streptomycetes nous
29	219	53.8	1996	18	AAW2607	Platenolide synthase
30	219	53.8	1996	18	AAW23717	Platenolide synthase
31	219	53.8	3729	18	AAW22603	Tylosone synthase
32	217	53.3	4881	21	AA823751	S. avermectilis ave
33	217	53.3	4881	22	AA65266	Streptomycetes averm
34	216	53.1	1891	19	AAW52844	Amycolatopsis medl
35	216	53.1	3413	19	AAW52849	A. mediterranei r1
36	215	52.8	5435	22	AAE10145	Streptomycetes nous
37	212	52.1	1688	19	AAW52848	A. mediterranei r1
38	212	52.1	2152	20	AA39298	SpnB a polyketide
39	212	52.1	2152	22	AA870966	S. spinosa protein
40	211	51.8	4630	21	AA77177	S. venezuelae vep
41	210	51.6	3170	20	AA39299	SpnC a polyketide
42	210	51.6	3170	22	AA870967	S. spinosa protein
43	209.5	51.5	186	22	AA819802	Polyketide synthase
44	209	51.4	7068	22	AAE10142	Streptomycetes nous
45	207.5	51.0	3201	22	AA882214	Polyketide synthase

ALIGNMENTS

RESULT 1
AA823749
ID AA823749 standard; Protein; 3972 AA.
XX
AC AA823749;
XX
DT 10-JAN-2001 (first entry)
XX
DE S. avermectilis avermectin aglycon synthase protein SEQ ID NO:3.
XX
KW Streptomycetes avermectilis; avermectin aglycon synthase; biosynthesis;
KW multifunctional enzyme; polyketide; avermectin; veterinary drug;
KW agrochemical.
XX
OS Streptomycetes avermectilis.
XX
PN NO200050605-A1.
XX
PD 31-AUG-2000.
XX
PF 23-FEB-2000; 2000MO-JP01041.
XX
PR 24-FEB-1999; 99JP-0046961.
XX
PA (KITA) KITASATO INST.
XX
PI Omura S, Ikeda H;
XX
DR WPI; 2000-565458/52.
XX
DR N-PSDB; AAA92301.
XX
PT Avermectin aglycone synthase DNA and proteins encoded by all or part of
it for the production of avermectin and its derivatives for drug and

```

PT agrochemical use -
XX
XX Claim 32; Page 203-223; 314pp; Japanese.
XX
XX The present sequence represents an avermectin aglycon synthase protein.
CC Also described are: (1) polypeptides encoded by all or part of the DNA;
CC (2) expression vectors containing the DNA; (3) host cells transformed by
CC the vectors; (4) preparation of the polypeptides by culture of the
CC transformants; (5) preparation of avermectin aglycon or its derivatives
CC by culture of transformed avermectin-producing microorganisms; and (6)
CC oligonucleotides of 5-60 bases in length containing sense or antisense
CC sequences from the avermectin aglycon synthase DNA. The enzymes are
CC useful for the production of modified forms of avermectin and of the
CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
CC and agrochemicals.
XX
XX Sequence 3972 AA;
XX
XX Query Match 100.0%; Score 407; DB 21; Length 3972;
XX Best Local Similarity 100.0%; Pred. No. 9,8e-46;
XX Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 LAHQTSADQRAALAEVLVDHYAAVLRHADPKAIAPDQSFRAIGPDSLTAVEFRNLIKAT 60
Db 1979 LAHQTSADQRAALAEVLVDHYAAVLRHADPKAIAPDQSFRAIGPDSLTAVEFRNLIKAT 2038
XX
XX 61 GLRLPVSLVFDHPPTPAKLAVH 82
XX
XX 2039 GLRLPVSLVFDHPPTPAKLAVH 2060
XX
XX RESULT 2
XX
XX AAG65264
XX ID AAG65264 standard; Protein; 3972 AA.
XX
XX AAG65264;
XX
XX 04-DEC-2001 (first entry)
XX
XX Streptomyces avermectilis protein SEQ ID NO: 4.
XX
XX Streptomyces avermectilis protein SEQ ID NO: 4.
XX
XX Avermectin aglycone synthase; AAs; avermectin derivative;
XX drug production; veterinary drug; pesticide.
XX
XX Streptomyces avermectilis.
XX
XX MO200162939-A1.
XX
XX 30-AUG-2001.
XX
XX 23-FEB-2001; 2001WO-JP01381.
XX
XX 24-FEB-2000; 2000JP-0047405.
XX
XX (KYOWA ) KYOWA HAKKO KOGYO KK.
XX (KITA ) KITASATO INST.
XX
XX Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
XX
XX MPI; 2001-582053/65.
XX N-PSDB; AAH79277.
XX
XX New modified avermectin aglycone synthase derived from Streptomyces
XX avermectilis used in production of 22,23-dihydroavermectin B1a used in
XX drugs and pesticides -
XX
XX Claim 4; Page 167-180; 257pp; Japanese.
XX
XX The present invention relates to the production of modified derivatives
XX of avermectin aglycone synthase (AAS) derived from Streptomyces
XX avermectilis. The activity of an acyl carrier protein (ACP),
XX beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
XX beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl

```

CC	reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC	suppressed. The process can be used in the production of drugs, veterinary
CC	drugs and pesticides. The present sequence is an S. avermitilis protein.
XX	
SO	Sequence 3972 AA;
	Query Match 100.0%; Score 407; DB 22; Length 3972;
	Best Local Similarity 100.0%; Pred. No. 9,8e-46;
	Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps
OY	1 LAHQTSDQRAALLELVRDHVAALVLRHADPKAIPDQSFRAIGFDSLTAVERNNLIKAT 60
DB	1979 LAHQTSDQRAALLELVRDHVAALVLRHADPKAIPDQSFRAIGFDSLTAVERNNLIKAT 20
OY	61 GLRLPVSLVFDHPPTPAKLAIVHL 82
DB	2039 GLRLPVSLVFDHPPTPAKLAIVHL 2060
RESULT 3	
ID	AA65268
XX	AA65268 standard; Protein; 3972 AA.
AC	AA65268;
XX	
XX	04-DEC-2001 (first entry)
DT	
DE	Streptomyces avermitilis protein derivative SEQ ID NO: 8.
XX	
XX	Avermectin aglycone synthase; AAS; avermectin derivative;
KW	drug production; veterinary drug; pesticide.
XX	
OS	Synthetic.
XX	
FN	MO200162939-A1.
PD	30-AUG-2001.
XX	
XX	23-FEB-2001; 2001MO-JF01381.
PF	
PR	24-FEB-2000; 2000JF-0047405.
XX	
PA	(KYOW) KYOWA HAKKO KOGYO KK.
XX	(KITA) KITASATO INST.
PI	Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
XX	
XX	WPI: 2001-582053/65.
DR	N-PSDB; AAH79279.
DR	
PT	New modified avermectin aglycone synthase derived from Streptomyces
XX	avermitilis used in production of 22,23-dihydroavermectin B1a used in
XX	drugs and pesticides -
XX	
XX	Claim 5; Page 235-248; 257pp; Japanese.
XX	
XX	The present invention relates to the production of modified derivatives
CC	of avermectin aglycone synthase (AAS) derived from Streptomyces
CC	avermitilis. The activity of an acyl carrier protein (ACP),
CC	beta-ketolacyl carrier protein synthase (KS), acyltransferase (AT),
CC	beta-ketolacyl carrier protein reductase (KR), dehydratase (DH), enoyl
CC	reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC	suppressed. The process can be used in the production of drugs, veterinary
CC	drugs and pesticides. The present sequence is derived from an S.
CC	avermitilis protein.
XX	
SO	Sequence 3972 AA;
	Query Match 100.0%; Score 407; DB 22; Length 3972;
	Best Local Similarity 100.0%; Pred. No. 9,8e-46;
	Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps

Db 1979 LAHQTSADQRAALIEVRDVAALRHADPKAIPDQSFALGFDLSLTAVEFRLIKAT 2038
 QY 61 GLRLPVSLVDFDHPPTPAKLAVHL 82
 Db 2039 GLRLPVSLVDFDHPPTPAKLAVHL 2060

RESULT 4

AAB3752
 ID AAB3752 standard; Protein; 5532 AA.

AC AAB3752;

DT 10-JAN-2001 (first entry)

DE 8. avermectin aglycon synthase protein SEQ ID NO:6.

XX Streptomyces avermectilis; avermectin aglycon synthase; biosynthesis;
 KW multifunctional enzyme; polyketide; avermectin; veterinary drug;
 KM agrochemical.

OS Streptomyces avermectilis.

PN WO200050605-A1.

PD 31-AUG-2000.

PF 23-FEB-2000; 2000WO-JP01041.

PR 24-FEB-1999; 99JP-0046961.

PA (KITA) KITASATO INST.

PI Omura S, Ikeda H;

DR WPI; 2000-565458/52.

DR N-PSDB; AAA92302.

PT Avermectin aglycone synthase DNA and proteins encoded by all or part of
 PT it for the production of avermectin and its derivatives for drug and
 PT agrochemical use -

PS Claim 32; Page 278-305; 314pp; Japanese.

XX The present sequence represents an avermectin aglycon synthase protein.
 CC Also described are: (1) polypeptides encoded by all or part of the DNA;
 CC (2) expression vectors containing the DNA; (3) host cells transformed by
 CC the vectors; (4) preparation of the polypeptides by culture of the
 CC transformants; (5) preparation of avermectin aglycon or its derivatives
 CC by culture of transformed avermectin-producing microorganisms; and (6)
 CC oligonucleotides of 5-60 bases in length containing sense or antisense
 CC sequences from the avermectin aglycon synthase DNA. The enzymes are
 CC useful for the production of modified forms of avermectin and of the
 CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
 CC and agrochemicals.

XX Sequence 5532 AA;

Query Match 59.7%; Score 243; DB 21; Length 5532;
 Best Local Similarity 59.8%; Pred. No. 4.8e-23;
 Matches 49; Conservative 5; Mismatches 28; Indels 0; Gaps 0;

QY 1 LAHQTSADQRAALIEVRDVAALRHADPKAIPDQSFALGFDLSLTAVEFRLIKAT 60
 Db 1756 LAQOTHEOQHHTTLALVRSHIATVLTGHTTPDTPDRAFDLGFDSLTAVALRRLSRTT 1815

QY 61 GLRLPVSLVDFDHPPTPAKLAVHL 82
 Db 1816 GLRLPTTLAFDHPPTTLTHHL 1837

RESULT 5

AAG65267
 ID AAG65267 standard; Protein; 5532 AA.

AC AAG65267;

DT 04-DEC-2001 (first entry)

DE Streptomyces avermectilis protein SEQ ID NO: 7.

XX Avermectin aglycone synthase; AAS; avermectin derivative;
 KW drug production; veterinary drug; pesticide.

OS Streptomyces avermectilis.

PN WO200162939-A1.

PD 30-AUG-2001.

PF 23-FEB-2001; 2001WO-JP01381.

PR 24-FEB-2000; 2000JP-0047405.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PA (KITA) KITASATO INST.

PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;

DR WPI; 2001-582053/65.

DR N-PSDB; AAH79278.

PT New modified avermectin aglycone synthase derived from Streptomyces
 PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
 PT drugs and pesticides -

PS Claim 4; Page 217-235; 257pp; Japanese.

XX The present invention relates to the production of modified derivatives
 CC of avermectin aglycone synthase (AAS) derived from Streptomyces
 CC avermectilis. The activity of an acyl carrier protein (ACP),
 CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
 CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is an S. avermectilis protein.

XX Sequence 5532 AA;

Query Match 59.7%; Score 243; DB 22; Length 5532;
 Best Local Similarity 59.8%; Pred. No. 4.8e-23;
 Matches 49; Conservative 5; Mismatches 28; Indels 0; Gaps 0;

QY 1 LAHQTSADQRAALIEVRDVAALRHADPKAIPDQSFALGFDLSLTAVEFRLIKAT 60
 Db 1756 LAQOTHEOQHHTTLALVRSHIATVLTGHTTPDTPDRAFDLGFDSLTAVALRRLSRTT 1815

QY 61 GLRLPVSLVDFDHPPTPAKLAVHL 82
 Db 1816 GLRLPTTLAFDHPPTTLTHHL 1837

RESULT 6

AAB31558
 ID AAB31558 standard; Protein; 6797 AA.

AC AAB31558;

DT 20-APR-2001 (first entry)

DE Pimaricin biosynthesis associated polyketide synthase polypeptide.

XX Polyketide synthase; oxidative modification; metabolite; antibiotic;
 KW anticancer; pimaricin.

```

OS Streptomyces natalensis.
XX
XX WO200077222-A1.
XX
XX 21-DEC-2000.
XX
XX 14-JUN-2000; 2000WO-EP06227.
XX
XX 14-JUN-1999; 99EP-0201893.
XX
XX (STRM ) DSM NV.
XX
XX Martin JF, Aparicio JF, Colina AJ;
XX
XX WPI; 2001-080693/09.
XX
XX N-PSDB; AAF24892.
XX
XX New polynucleotides encoding enzymes involved in the biosynthesis of
XX pimaricin, useful for modifying the biosynthesis of pimaricin and in
XX the synthesis of new compounds -
XX
XX Disclosure; Page 81-101; 116pp; English.
XX
XX The present sequence represents a polyketide synthase which is associated
XX with the biosynthesis of pimaricin. The polyketide synthase polypeptide
XX is useful for the oxidative modification of a methyl group of a suitable
XX compound, e.g. a bioactive compound including a secondary metabolite,
XX antibiotics and anticancer agents. Recombinant cells comprising the
XX gene are useful for the production of pimaricin. The polyketide synthase
XX polynucleotide may be over expressed in Streptomyces, leading to an
XX increase in the biosynthesis of pimaricin, as a source of primers for
XX amplification reaction and as probes.
XX
XX Sequence 6797 AA;
XX
XX Query Match 59.0%; Score 240; DB 22; Length 6797;
XX Best Local Similarity 59.8%; Pred. No. 1.6e-22;
XX Matches 49; Conservative 14; Mismatches 19; Indels 0; Gaps 0;
XX
XX 1 LAHQTSADQRAALLELVDRHVAALVRHADPKAIAPDQSPFALGFDLSLTAVEFRNLTKAT 60
XX |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
XX LAGISTIERREBALDDVRAQIATVLRGHAGPRTIAPDRAFQDLGDSLTALERNLTKAT 4952
XX
XX 61 GLRLPVSIVDPDHPPTPAKLAHNL 82
XX |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
XX GLRLPATTVPDYPVDALAHNL 4974
XX
XX Db 4953 GLRLPATTVPDYPVDALAHNL 4974
XX
XX RESULT 7
XX AAB23750
XX ID AAB23750 standard; Protein; 6239 AA.
XX
XX AC AAB23750;
XX
XX 10-JAN-2001 (first entry)
XX
XX S. avermitilis avermectin aglycon synthase protein SEQ ID NO:4.
XX
XX Streptomyces avermitilis; avermectin aglycon synthase; biosynthesis;
XX multifunctional enzyme; polyketide; avermectin; veterinary drug;
XX agrochemical.
XX
XX Streptomyces avermitilis.
XX
XX WO200050605-A1.
XX
XX 31-AUG-2000.
XX
XX 23-FEB-2000; 2000WO-JP01041.
XX
XX 24-FEB-1999; 99JP-0046961.
XX
XX (KITA ) KITASATO INST.
XX

```

```

XX Omura S, Ikeda H;
XX
XX WPI; 2000-565458/52.
XX
XX N-PSDB; AAA92301.
XX
XX Avermectin aglycone synthase DNA and proteins encoded by all or part of
XX it for the production of avermectin and its derivatives for drug and
XX agrochemical use -
XX
XX Claim 32; Page 223-254; 314pp; Japanese.
XX
XX The present sequence represents an avermectin aglycon synthase protein.
XX Also described are: (1) polypeptides encoded by all or part of the DNA;
XX (2) expression vectors containing the DNA; (3) host cells transformed by
XX the vectors; (4) preparation of the polypeptides by culture of the
XX transformants; (5) preparation of avermectin aglycon or its derivatives
XX by culture of transformed avermectin-producing microorganisms; and (6)
XX oligonucleotides of 5-60 bases in length containing sense or antisense
XX sequences from the avermectin aglycon synthase DNA. The enzymes are
XX useful for the production of modified forms of avermectin and of the
XX intermediates in its biosynthesis, for use as drugs, veterinary drugs
XX and agrochemicals.
XX
XX Sequence 6239 AA;
XX
XX Query Match 57.2%; Score 233; DB 21; Length 6239;
XX Best Local Similarity 61.0%; Pred. No. 1.3e-21;
XX Matches 50; Conservative 8; Mismatches 24; Indels 0; Gaps 0;
XX
XX 1 LAHQTSADQRAALLELVDRHVAALVRHADPKAIAPDQSPFALGFDLSLTAVEFRNLTKAT 60
XX |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
XX LARQSAGERRQALLRLVRSVVAALVLRGHAGPRTIAPDRAFQDLGDSLTALERNLTKAT 6091
XX
XX 61 GLRLPVSIVDPDHPPTPAKLAHNL 82
XX |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
XX GLRLPATTVPDYPVDALAHNL 6113
XX
XX Db 6092 GLRLPATTVPDYPVDALAHNL 6113
XX
XX RESULT 8
XX AAG65265
XX ID AAG65265 standard; Protein; 6239 AA.
XX
XX AC AAG65265;
XX
XX 04-DEC-2001 (first entry)
XX
XX Streptomyces avermitilis protein SEQ ID NO: 5.
XX
XX Avermectin aglycone synthase; AAS; avermectin derivative;
XX drug production; veterinary drug; pesticide.
XX
XX Streptomyces avermitilis.
XX
XX WO200162939-A1.
XX
XX 30-AUG-2001.
XX
XX 23-FEB-2001; 2001WO-JP01381.
XX
XX 24-FEB-2000; 2000JP-0047405.
XX
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX (KITA ) KITASATO INST.
XX
XX Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
XX
XX WPI; 2001-582053/65.
XX
XX N-PSDB; AAH79277.
XX
XX New modified avermectin aglycone synthase derived from Streptomyces
XX avermitilis used in production of 22,23-dihydroavermectin B1a used in
XX drugs and pesticides -
XX

```

XX Claim 4; Page 180-201; 257p; Japanese.
 PS The present invention relates to the production of modified derivatives
 XX of avermectin aglycone synthase (AAS) derived from Streptomyces
 CC avermectilis. The activity of an acyl carrier protein (ACP),
 CC beta-ketacyl carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketacyl carrier protein reductase (KR), dehydratase (DH), enoyl
 CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is an S. avermectilis protein.
 XX
 SQ Sequence 6239 AA;
 Query Match 57.2%; Score 233; DB 22; Length 6239;
 Best Local Similarity 61.0%; Pred. No. 1,38-21;
 Matches 50; Conservative 8; Mismatches 24; Indels 0; Gaps 0;
 QY 1 LAHQTSADQRAALLELVHDHVAALRHADPKAIPDQSPFALGDSITTAVERNLIKAT 60
 DB 6032 LARGSAGERQALRLVRSVHAVALGHSADGIDASRAFRGLTAVELRNRLTAAT 6091
 QY 61 GLRLPVSLVDFHPTPAKLAVHL 82
 DB 6092 GLRLRATLAFDPTPALAEHL 6113
 RESULT 9
 AAB18638
 ID AAB18638 standard; Protein; 3739 AA.
 AC AAB18638;
 XX
 DT 22-JUN-2001 (first entry)
 DE Antino acid sequence of narbonolide synthase subunit 2 (PICAITI).
 XX
 KM Narbonolide synthase; polyketide synthase gene; narbonolide polyketide;
 KM antibiotic; C12-hydroxylase; pick; desosamine biosynthesis;
 KM desosaminyl transferase enzyme; ketolide; beta-glucosidase enzyme;
 KM picromycin biosynthesis.
 XX
 OS Streptomyces venezuelae.
 XX
 PN US6117659-A.
 PD 12-SEP-2000.
 XX
 PF 27-MAY-1999; 99US-0320878.
 XX
 PR 28-MAY-1998; 98US-0087080.
 PR 22-SEP-1998; 98US-0100880.
 PR 08-FEB-1999; 99US-0119139.
 PR 20-MAY-1999; 99US-0134990.
 PR 30-APR-1997; 97US-0846247.
 PR 06-MAY-1998; 98US-0073538.
 PR 28-AUG-1998; 98US-0141908.
 XX
 PA (KOSA-) KOSAN BIOSCIENCES INC.
 PI Ashley G, Beclach MC, Beclach M, Tang L, McDaniel R;
 XX
 DR WPI; 2000-610844/58.
 PT New recombinant pick hydroxylase gene of Streptomyces venezuelae useful
 PT for converting ketolide to antibiotics and as antibiotics and
 PT intermediates in the synthesis of compounds with pharmaceutical value
 XX
 PS Disclosure; Columns 11-12; 117p; English.
 XX The present sequence represents a narbonolide synthase subunit 2
 CC (PICAITI). The nucleotide sequence encoding it is used in the course of

CC the invention. The specification describes a recombinant DNA compound
 CC expressing recombinant polyketide synthase genes in host cells for the
 CC production of narbonolide, narbonolide derivatives and polyketides that
 CC are useful as antibiotics and as intermediates in the synthesis of
 CC compounds with pharmaceutical value. The DNA compounds may also encode
 CC a C12-hydroxylase (pick), desosamine biosynthesis and desosaminyl
 CC transferase enzymes (useful for conversion of ketolides to antibiotics),
 CC and the beta-glucosidase enzyme (involved in picromycin biosynthesis).
 CC These compounds are also useful for increasing the antibiotic activity
 CC of a compound relative to the unhydroxylated compound. The recombinant
 CC host cells are useful as genetic systems that allow rapid engineering
 CC of the narbonolide polyketide synthase. These would be valuable for
 CC creating novel ketolide analogs for pharmaceutical applications.
 XX
 SQ Sequence 3739 AA;
 Query Match 56.3%; Score 229; DB 21; Length 3739;
 Best Local Similarity 59.8%; Pred. No. 2,48-21;
 Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;
 QY 1 LAHQTSADQRAALLELVHDHVAALRHADPKAIPDQSPFALGDSITTAVERNLIKAT 60
 DB 3560 LAAMTPDDRVAHLRLDLVRTHVATVGLGHPRSVDLEBARFDTGDSITTAVERNRLTAAT 3619
 QY 61 GLRLPVSLVDFHPTPAKLAVHL 82
 DB 3620 GLRLPATLVDFDPTPELAGHL 3641
 RESULT 10
 AAY77193
 ID AAY77193 standard; Protein; 3739 AA.
 AC AAY77193;
 XX
 DT 05-JUN-2000 (first entry)
 DE S. venezuelae macrolide biosynthetic enzyme pikaiti, SEQ ID NO:33.
 XX
 KM Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
 KM nomethymycin; natomycin; polyhydroxyalkanoate monomer synthase;
 KM biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;
 KM chronic obstructive pulmonary disease; respiratory inflammation;
 KM hypercholesterolemia; crop protection agent.
 XX
 OS Streptomyces venezuelae ATCC15439.
 XX
 PN WO200000620-A2.
 PD 06-JAN-2000.
 XX
 PF 25-JUN-1999; 99WO-US14398.
 XX
 PR 26-JUN-1998; 98US-0105537.
 XX
 PA (MINU) UNIV MINNESOTA.
 PI Sherman DH, Liu H, Xue Y, Zhao L;
 XX
 DR WPI; 2000-160679/14.
 DR N-PDB; AA287298.
 PT Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.,
 PT synthesis of methymycin and pikromycin -
 XX
 PS Claim 19; Page 403-415; 438p; English.
 XX
 CC The invention relates to an isolated and purified nucleic acid segment
 CC comprising a desosamine biosynthetic gene cluster, a fragment or its
 CC biologically active variant, where the nucleic acid sequence is not
 CC derived from the eryc gene cluster of Saccharopolyspora erythraea or
 CC Streptomyces antibioticus. The invention also relates to a macrolide
 CC biosynthetic gene cluster, or fragments thereof. The macrolide

XX		Disclosure; Figure 31; 438bp; English.
CC	The invention relates to an isolated and purified nucleic acid segment	
CC	comprising a deoxamine biosynthetic gene cluster, a fragment or its	
CC	biologically active variant, where the nucleic acid sequence is not	
CC	derived from the eryC gene cluster of Saccharopolyspora erythraea or	
CC	Streptomyces antibioticus. The invention also relates to a macrolide	
CC	biosynthetic gene cluster, or fragments thereof. The macrolide	
CC	biosynthetic gene cluster encodes proteins which synthesise methymycin,	
CC	pikromycin, nemethymycin, narbomycin or a combination of these	
CC	compounds. Recombinant or augmented cells comprising the deoxamine	
CC	and/or macrolide biosynthetic gene clusters are useful for the	
CC	production of biologically active macroclides. The macroide biosynthetic	
CC	proteins are useful for synthesis of methymycin, pikromycin,	
CC	nemethymycin and narbomycin. The alternative termination of polyketide	
CC	synthesis may be useful to prepare novel antibiotics and	
CC	polyhydroxyalkanoate (PHA) monomers. The compounds produced by the	
CC	recombinant host cells are useful as biopolymers, e.g., in packaging or	
CC	biomedical applications, to engineer PHA monomer synthases or to prepare	
CC	biologically active agents, such as chemotherapeutics,	
CC	immunosuppressants, agents to treat asthma, chronic obstructive pulmonary	
CC	disease as well as other diseases involving respiratory inflammation,	
CC	cholesterol-lowering agents or macrolide-based antibiotics which are	
CC	active against a variety of organisms, e.g., bacteria, including	
CC	multi-drug resistant pneumococci and other respiratory pathogens, as well	
CC	as viral parasitic pathogens, or as crop protection agents (e.g.,	
CC	fungicides or insecticides) via expression of polyketides in plants.	
CC	Sequences AAY77200-Y77203 and AAY8097 represent proteins encoded by the	
CC	macroide biosynthetic gene cluster (pk) from Streptomyces venezuelae	
CC	ATCC 15439, as given in figure 31.	
XX		
SQ	Sequence 3739 AA;	
Query Match	56.3%; Score 229; DB 21; Length 3739;	
Best Local Similarity	59.8%; Pred. No. 2,4e-21;	
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;		
QY	1 LAHOTSADQPAALLBELVRDHYAAVLRNHPDKAIAPDOSFALGFDSTAVEFNRLIKAT 60	
DY	LAAMTPDDRRVAHRIDIVRTAVATVLGHGPSRVDLERAFRDGTGDSLTAVELNRINMAAT 3619	
QY	61 GLRLPVSILVFDPHTPAKLAVAL 82	
DY	3620 GLRLPATLVFPDHPTEGLAGHL 3641	
RESULT 12		
AAY67202		
ID	AAY67202 standard; protein: 3739 AA.	
XX		
AC	AAY67202;	
XX		
DT	23-MAR-2000 (first entry)	
DE	Narbonolide synthase subunit 2 (PICII) protein sequence.	
XX		
KM	Narbonolide polyketide synthase; PKS; narbonolide synthase subunit 2;	
PICII; antibiotic production; narbomycin; pikromycin; ketolide.		
OS	Streptomycetes venezuelae.	
XX		
PN	WO961599-A2.	
PD	02-DEC-1999.	
PF	27-MAY-1999; 99MO-US11814.	
XX		
PR	28-MAY-1998; 98US-0087080.	
PR	28-AUG-1998; 98US-0141908.	
PR	22-SEP-1998; 98US-0100880.	
PR	08-FEB-1999; 99US-0119139.	
XX		

XX Claim 15; Page 255-260; 266pp; English.
XX
CC The present invention relates to the cloning and sequencing of the gene
CC cluster encoding a modular type I polyketide synthase (PKS) enzyme
CC involved in the biosynthesis of the macrocyclic antibiotic nystatin.
CC The nystatin PKS is useful as antifungal antibiotics. The present
CC sequence is a PKS type I encoding Streptomyces noursei
CC nystatin gene, NysI (complete).
XX
SQ Sequence 9477 AA;
Query Match 56.3%; Score 229; DB 22; Length 9477;
Best Local Similarity 59.8%; Pred. No. 8.2e-21;
Matches 49; Conservative 5; Mismatches 28; Indels 0; Gaps 0;
QY 1 LAHQTADQRAALLEVRDHAVALRHADPKAIPDQSFALGFDSTLAVEFNNLIKAT 60
DB 7801 LAEQPADHDHAILALVTEKAAVIGHADHDLEPDPFRDLGFDSTLAVDLRNQLTAET 7860
QY 61 GLRLPVSILVDPDHPPTPAKLAVHL 82
DB 7861 GLRLPALTVPDHPNPAALAH 7882
RESULT 14
AAV77180
ID AAV77180 standard; Protein; 12199 AA.
XX
AC AAV77180;
XX
DT 05-JUN-2000 (first entry)
XX
DE S. venezuelae plk gene cluster-encoded protein, SEQ ID NO:6.
XX
KM Desosamine biosynthesis; macrocyclic; polyketide; methymycin; pikromycin;
KM neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
KM biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;
KM chronic obstructive pulmonary disease; respiratory inflammation;
KM hypercholesterolemia; crop protection agent.
XX
OS Streptomyces venezuelae ATCC15439.
XX
PN WO20000620-A2.
PD 06-JAN-2000.
XX
PF 25-JUN-1999; 99WO-US14398.
XX
PR 26-JUN-1998; 98US-0105537.
XX
PA (MINU) UNIV MINNESOTA.
PI Sherman DH, Liu H, Xue Y, Zhao L;
XX
XX WPI; 2000-160679/14.
DR N-PSDB; AAZ87285.
XX
PT Desosamine and macrocyclic biosynthetic gene clusters, useful for, e.g.,
PT synthesis of methymycin and pikromycin -
XX
PS Claim 19; Page 315-353; 438pp; English.
XX
CC The invention relates to an isolated and purified nucleic acid segment
CC comprising a desosamine biosynthetic gene cluster, a fragment or its
CC biologically active variant, where the nucleic acid sequence is not
CC derived from the eryC gene cluster of Saccharopolyspora erythraea or
CC Streptomyces antibioticus. The invention also relates to a macrocyclic
CC biosynthetic gene cluster, or fragments thereof. The macrocyclic
CC biosynthetic gene cluster encodes proteins which synthesize methymycin,
CC pikromycin, neomethymycin, narbomycin or a combination of these
CC compounds. Recombinant or augmented cells comprising the desosamine
CC and/or macrocyclic biosynthetic gene clusters are useful for the production

CC of biologically active macrocyclics. The macrocyclic biosynthetic proteins
CC are useful for synthesis of methymycin, pikromycin, neomethymycin and
CC narbomycin. The alternative termination of polyketide synthase may be
CC useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA)
CC monomers. The compounds produced by the recombinant host cells are useful
CC as biopolymers, e.g., in packaging or biomedical applications, to
CC engineer PHA monomer synthases or to prepare biologically active agents,
CC such as chemotherapeutics, immunosuppressants, agents to treat asthma,
CC chronic obstructive pulmonary disease as well as other diseases involving
CC respiratory inflammation, cholesterol-lowering agents or macrocyclic-based
CC antibiotics which are active against a variety of organisms, e.g.,
CC bacteria, including multi-drug resistant pneumococci and other
CC respiratory pathogens, as well as viral parasitic pathogens, or as crop
CC protection agents (e.g., fungicides or insecticides) via expression of
CC polyketides in plants. The present sequence represents a protein
CC encoded by the macrocyclic biosynthetic gene cluster (plk) from
CC Streptomyces venezuelae ATCC 15439.
XX
SQ Sequence 12199 AA;
Query Match 56.3%; Score 229; DB 21; Length 12199;
Best Local Similarity 59.8%; Pred. No. 1.2e-20;
Matches 49; Conservative 8; Mismatches 25; Indels 0; Gaps 0;
QY 1 LAHQTADQRAALLEVRDHAVALRHADPKAIPDQSFALGFDSTLAVEFNNLIKAT 60
DB 8831 LAAMTPDDRAHARDLVRTHTVATVLCGTPSRVDLERARFDGSLTAVELNRLNMAAT 8890
QY 61 GLRLPVSILVDPDHPPTPAKLAVHL 82
DB 8891 GLRLPALTVPDHPPELAGHL 8912
RESULT 15
AAV67203
ID AAV67203 standard; protein; 1560 AA.
XX
AC AAV67203;
XX
DT 23-MAR-2000 (first entry)
XX
DE Narbonolide synthase subunit 3 (PICAIIT) protein sequence.
XX
KM Narbonolide polyketide synthase; PKS; narbonolide synthase subunit 3;
KM PICAIIT; antibiotic production; narbomycin; picromycin; ketolide.
XX
OS Streptomyces venezuelae.
XX
PN WO961599-A2.
PD 02-DEC-1999.
XX
PF 27-MAY-1999; 99WO-US11814.
XX
PR 28-MAY-1998; 98US-0087080.
PR 28-AUG-1998; 98US-0141908.
PR 22-SEP-1998; 98US-0100880.
PR 08-FEB-1999; 99US-0119139.
XX
PA (KOSA-) KOSAN BIOSCIENCES INC.
PI Ahlley G, Betlach MC, Betlach M, McDaniel R, Tang L;
XX
XX WPI; 2000-072618/06.
DR N-PSDB; AAZ56001.
XX
PT New recombinant DNA encoding a domain of narbonolide polyketide
PT synthase, for production of ketolide antibiotics -
XX
PS Example 2; Page 14; 98pp; English.
XX
XX This is the Streptomyces venezuelae narbonolide synthase subunit 3,
CC PICAIIT protein sequence. The invention relates to recombinant DNA

CC containing a coding sequence for a narbonolide polyketide synthase
 CC (PKS). Polyketides are compounds synthesised from 2-carbon units through
 CC a series of condensations and subsequent modifications. Modular PKSs are
 CC responsible for the production of many antibiotics including picromycin.
 CC The narbonolide PKS consists of a loading module, six extender modules,
 CC and two thioester domains. Four proteins make up the narbonolide PKS
 CC (PICAI, PICAI1, PICAI11 and PICAIIV). PICAI includes the loading module
 CC and extender modules 1 and 2, PICAI1 includes extender modules 3 and 4,
 CC PICAI11 includes extender module 5 and PICAIIV includes extender module 6
 CC and a type II thioesterase domain. The second type II thioesterase
 CC domain is found on the PICB protein. The nucleotide sequences encoding
 CC all of these proteins can be isolated in recombinant form from the
 CC recombinant cosmid pKOS023-27 (see AAZ56001). Narbonolide is
 CC desosaminylated in *S. venezuelae* to yield narbomycin, the desosaminyl
 CC transferase enzyme is required for this conversion, and the desosamine
 CC biosynthetic genes are also found in cosmid pKOS023-27. The recombinant
 CC DNA of the invention is used to express, in transformed cells,
 CC narbonolide (or its derivatives) or other ketolides (particularly
 CC hybrids), which may then be converted (e.g. by other enzymes
 CC recombinantly expressed in the same hosts) to polyketide antibiotics or
 CC their intermediates. The antibiotics are useful in human or veterinary
 CC medicine.

XX Sequence 1560 AA;

Query Match 55.3%; Score 225; DB 21; Length 1560;

Best Local Similarity 60.0%; Pred. No. 2,66-21;

Matches 45; Conservative 8; Mismatches 22; Indels 0; Gaps 0;

QY 8 DQRAALLIELVNVAALVLRHADPKAIAPDQSFRLGFDLSLTAVERNLIKATGLRLPVS 67

DB 1400 ERRPALTLVTRTHAAVLGHSSPDRAVAPRAFTLGFDSLTAVQLRNQLSTVVGNRLPAT 1459

QY 68 LVPDHPTPAKLAAYL 82

DB 1460 TVPDHPTPALAAHL 1474

Search completed: June 17, 2003, 13:02:31
 Job time : 5.01696 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 12:53:02 ; Search time 8.22676 Seconds
(without alignments)
4458.180 Million cell updates/sec

Title: US-09-914-286-3_COPY_1715_1892

Perfect score: 953
Sequence: 1 GTTLITGCTGALATHLTHL.....LAHHRTHLPATSIAMGTW 178

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing filter 45 summaries

Database :

1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phase:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_rvtrius:*
17: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	953	100.0	3972	2 Q9S0R8	Q9S0R8 streptomyc
2	948	99.5	6239	2 Q9S0R7	Q9S0R7 streptomyc
3	894.5	93.9	3626	2 Q9F779	Q9F779 streptomyc
4	573	60.1	3970	2 Q93H18	Q93H18 streptomyc
5	501	52.6	6315	2 Q9ADU6	Q9ADU6 polyanthum
6	500	52.5	5435	2 Q9L4X2	Q9L4X2 streptomyc
7	483	50.7	11096	2 Q9L4X3	Q9L4X3 streptomyc
8	481	50.5	5644	2 Q93N18	Q93N18 streptomyc
9	480	50.4	1879	2 Q30768	Q30768 streptomyc
10	479	50.3	1569	2 Q30767	Q30767 streptomyc
11	475	49.8	2066	2 Q9L4X1	Q9L4X1 streptomyc
12	471.5	49.5	9507	2 Q9EWA1	Q9EWA1 streptomyc
13	471	49.4	2034	2 Q93N17	Q93N17 streptomyc
14	471	49.4	3576	16 Q9EX53	Q9EX53 streptomyc
15	471	49.4	4840	2 Q93H04	Q93H04 streptomyc
16	468	49.1	1808	2 Q9EWA2	Q9EWA2 streptomyc

17	467	49.0	2731	2 Q923T9	Q923T9 pseudomonas
18	464	48.7	8817	2 Q53840	Q53840 polyanthum
19	461.5	48.4	9477	2 Q9L4X3	Q9L4X3 streptomyc
20	457	48.0	1864	2 Q33955	Q33955 streptomyc
21	457	47.9	10917	2 Q93N16	Q93N16 streptomyc
22	456.5	47.9	6797	2 Q9X393	Q9X393 streptomyc
23	454	47.6	3613	2 Q93HJ1	Q93HJ1 streptomyc
24	452	47.4	4881	2 Q9S0R3	Q9S0R3 streptomyc
25	452	47.4	6260	2 Q54299	Q54299 streptomyc
26	449	47.1	8563	2 Q54297	Q54297 streptomyc
27	447	46.9	3654	2 Q30766	Q30766 streptomyc
28	445.5	46.7	9510	2 Q93N19	Q93N19 streptomyc
29	445	46.7	2066	2 Q923T8	Q923T8 pseudomonas
30	444.5	46.6	6146	2 Q93HJ5	Q93HJ5 streptomyc
31	442	46.4	4340	2 Q30764	Q30764 streptomyc
32	440	46.2	10223	2 Q54296	Q54296 streptomyc
33	439	46.1	4735	2 Q54666	Q54666 amycolatops
34	436	45.8	3939	2 Q93HJ3	Q93HJ3 streptomyc
35	436	45.8	5588	2 Q9ALM2	Q9ALM2 saccharopol
36	436	45.8	7525	2 Q9KIE0	Q9KIE0 streptomyc
37	435	45.6	4498	2 Q93HE5	Q93HE5 streptomyc
38	431	45.2	4151	16 Q53490	Q53490 mycobacteri
39	429.5	45.1	6420	2 P95814	P95814 streptomyc
40	428.5	45.0	4613	2 Q92G15	Q92G15 streptomyc
41	427	44.8	4290	2 Q9MXC0	Q9MXC0 microcomosp
42	425	44.6	3192	2 Q9L4M4	Q9L4M4 streptomyc
43	424.5	44.5	1191	2 Q9RHV0	Q9RHV0 streptomyc
44	423	44.4	4472	2 Q33954	Q33954 streptomyc
45	422	44.3	1907	2 Q93H19	Q93H19 streptomyc

ALIGNMENTS

RESULT 1

Q9S0R8 PRELIMINARY; PRT; 3972 AA.
AC Q9S0R8;
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 1.
GN AVEA1.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9380548; PubMed=10449723;
RA Ikeda H., Nonomiyama T., Usami M., Ohta T., Omura S.,
RT "Organization of the biosynthetic gene cluster for the polyketide
antihelminthic macrocyclic avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL; AB032367; BAA84474.1; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantane-attach.
DR InterPro; IPR000834; Zn-carboxypept.
DR Pfam; PF00698; Acyl-transf. 3.
DR Pfam; PF00109; Ketoacyl-synt. 2.
DR Pfam; PF02801; ketoacyl-synt. C; 2.
DR Pfam; PF00550; pp-binding; 3_
DR PROSITE; P550075; ACP-DOMAIN; 3.
DR PROSITE; P500606; B-KETOACYL-SYNTHASE; 2.
DR PROSITE; P500133; CARBOXYPEPT_ZN_2; UNKNOWN_2.
DR PROSITE; P500012; PHOSPHOPANTETHEINE; 2.
KM Phosphopantetheine; Transferase.
SQ SEQUENCE 3972 AA; 416852 MW; 2A2936958032B1C3 CRC64;

Query Match 100.0%; Score 953; DB 2; Length 3972;
Best Local Similarity 100.0%; Pred. No. 7.1e-77;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy		1	GTTTLTGCGALATLTHNLTTTHOFTONLLSTRGPHTPHANOHTTLOOKGHLITTT	60
Dd		1715	GTTTLTGCGALATLTHNLTTTHOFTONLLSTRGPHTPHANOHTTLOOKGHLITTT	1774
Oy		61	CDTSNPDLQOLLNTPPOHPLTIVHTAGILDDATLTNLPTQOLNNVLRKAKSAHLH	120
Dd		1775	CDTSNPDLQOLLNTPPOHPLTIVHTAGILDDATLTNLPTQOLNNVLRKAKSAHLH	183
Oy		121	QLTOHTPLTAFLYSSAAATFGAPQANYYAANAAYLDALAHHRHTHHLPATSIAMGTW	178
Dd		1835	QLTOHTPLTAFLYSSAAATFGAPQANYYAANAAYLDALAHHRHTHHLPATSIAMGTW	1892
 RESULT 2 O9SOR7				
ID	O9SOR7	PRELIMINARY;	PRT; 6239 AA.	
AC	O9SOR7;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Type I polyketide synthase AVES 2.			
GN	AVB2.			
OS	Streptomyces avermitilis.			
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
OX	Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.			
NCBI	_Taxid=33903;			
RN	[1]			
RA	SEQUENCE FROM N.A.			
RX	MEDLINE=99380548; PubMed=10449723;			
IKeda H., Nomomiya T., Usami M., Ohta T., Omura S.;				
"Organization of the biosynthetic gene cluster for the polyketide				
antelmintic macrocyclic avermectin in Streptomyces avermitilis.";				
Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).				
DR	EMBL; AB032367; BAA84475.1.			
DR	InterPro; IPRO01227; Ac_transferase.			
DR	InterPro; IPRO00794; Ketoacyl-synt.			
DR	InterPro; IPRO03880; Pantane attach.			
DR	InterPro; IPRO00169; SHprot acite.			
DR	InterPro; IPRO00834; Zn cardoSept.			
DR	Pfam; PF00698; ACRYL_transf. 4.			
DR	Pfam; PRO0109; ketoacyl-synt. 4.			
DR	Pfam; PF02801; ketoacyl-synt-Ct. 4.			
DR	Pfam; PF00550; PD-binding; 4.			
DR	PROSITE; PS50075; ACP DOMAIN; 4.			
DR	PROSITE; PS00606; B_KETOACYL_SYNTHASE; 4.			
DR	PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_4.			
DR	PROSITE; PS00012; PHOSPHOPANTHEINE; 3.			
DR	PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.			
KW	Phosphopantetheine; Transferase.			
SC	SEQUENCE 6239 AA; 666283 MW; A706280940B502FA CRC64;			
 Query Match 99.5%; Score 948; DB 2; Length 6239; Best Local Similarity 99.4%; Pred. No. 3.4e-76; Matches 177; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
Oy		1	GTTTLTGCGALATLTHNLTTTHOFTONLLSTRGPHTPHANOHTTLOOKGHLITTT	60
Dd		2306	GTTTLTGCGALATLTHNLTTTHOFTONLLSTRGPHTPHANOHTTLOOKGHLITTT	2365
Oy		61	CDTSNPDLQOLLNTPPOHPLTIVHTAGILDDATLTNLPTQOLNNVLRKAKSAHLH	120
Dd		2366	CDTSNPDLQOLLNTPPOHPLTIVHTAGILDDATLTNLPTQOLNNVLRKAKSAHLH	2425
Oy		121	QLTOHTPLTAFLYSSAAATFGAPQANYYAANAAYLDALAHHRHTHHLPATSIAMGTW	178
Dd		2426	QLTOHTPLTAFLYSSAAATFGAPQANYYAANAAYLDALAHHRHTHHLPATSIAMGTW	2483
 RESULT 3 O9F7T9				
AC	O9F7T9	PRELIMINARY;	PRT; 3626 AA.	
KC	O9F7T9;			

DT	01-MAR-2001 (TREMBLrel. 16, Created)
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DS	Avermectin polyketide synthase (Fragment).
OS	Streptomyces avermitilis.
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
CC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX	NCBI_TaxID=33903;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=ATCC31271;
RA	Hong Y.-S., Lee J.J.;
RT	"Targeted Gene Disruption of the Avermectin O-methyltransferase gene
RL	and polyketide synthase gene from Streptomyces avermitilis.";
RU	Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AF275943; AAG09812.1;
DR	InterPro; IPR001227; Ac transferase.
DR	InterPro; IPR000794; Ketoacyl-synt.
DR	InterPro; IPR003880; Pantane attach.
DR	Pfam; PF00698; Acyl_transf_3
DR	Pfam; PF00109; ketoacyl-synt.; 2.
DR	Pfam; PF02801; ketoacyl-synt.C; 2.
DR	Pfam; PF00550; pp-binding; 2.
DR	PROSITE; PS50075; ACP DOMAIN; 2.
DR	PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
FT	PROSITE; PS00012; PHOSPHOPANTHETINE; UNKNOWN_1.
FT	NON TER 3626 3626
SQ	SEQUENCE 3626 AA; 380557 MW; 6272F5F088C1A8D0 CRC64;
Query Match	93.9%; Score 894.5; DB 2; Length 3626;
Best Local Similarity	94.4%; Pred. No. 1,2e-71;
Matches 169; Conservative 1; Mismatches 8; Indels 1; Gaps 1;	
Oy	1 GTTITIGTGALATHLTHTLTTHTOPTHQLLLTSRTGHPHQAHLTLLOOKGIHLITTT 60
Dd	1681 GTTITIGTGALATHLTHTLTTHTOPTHQLLLTSRTGHPHQAHLTLLOOKGIHLITTT 1740
Oy	61 CDTSPNPOL-QQLINTIPPOHEPTVHTKGIIDDLTLTLTPQLNNVLRAKHSALL 119
Dd	1741 CDTSPRPNHNNSLNTLPQHPVTVTHTGILDDATLTNLTPQLNNVLRAKHSALL 1800
Oy	120 HOLTOHPLTFNVLYSSAAATFGAPGOQANAAAYADALAHNHHTHLPTSIAWGTV 178
Dd	1801 HOLTOHPLTRFNLYSSAAATFGAPGOQANAAAYADALAHNRHTHLPTSIAWGTV 1859
RESULT 4	
ID	O93HI8 PRELIMINARY; PRT; 3970 AA.
AC	O93HI8.
DT	01-DEC-2001 (TREMBLrel. 19, Created)
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE	Modular polyketide synthase.
OS	OLM4.
OS	Streptomyces avermitilis.
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
CC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX	NCBI_TaxID=33903;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	MEDLINE=21477403; PubMed=11572948;
RA	Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA	Shinohe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA	Kikuchi H., Shiba T., Sakai Y., Hartori M.;
RT	"Genome sequence of an industrial microorganism Streptomyces
RL	avermitilis: Deducing the ability of producing secondary
RU	metabolites";
RU	Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR	EMBL; AB070940; BAB69199.1;
DR	InterPro; IPR001227; Ac transferase.
DR	InterPro; IPR002085; Adn_zn_family.
DR	InterPro; IPR000794; Ketoacyl-synt.

DR InterPro: IPR003880; Ppantne_attach.
 DR InterPro: IPR002364; OOR_zeta_crystal.
 DR Pfam: PF00698; Acyl_transferase; 2.
 DR Pfam: PF00107; adh_zinc; 1.
 DR Pfam: PF00109; ketoacyl-synt; 2.
 DR Pfam: PF02801; ketoacyl-synt_C; 2.
 DR Pfam: PF00550; pp-binding; 2.
 DR PROSITE: PS50075; ACP_DOMAIN; 2.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 DR PROSITE: PS00162; OOR_ZETA_CRYSTAL; UNKNOWN_1.
 KW Phosphopantetheine.
 SQ SEQUENCE 3970 AA; 412657 MW; A93E583FAFA66C47 CRC64;

Query Match 60.1%; Score 573; DB 2; Length 3970;
 Best local Similarity 62.9%; Pred. No. 1,4e-42;

Matches 112; Conservative 12; Mismatches 54; Indels 0; Gaps 0;

QY 1 GTTLTGGTGAATLTHLTHLTHQPTQHLLTSRTGPHRPHAOHLTTQLOQKIHLLTTT 60
 DB 1458 GTALITGGTGLGALLARHLVTRHGVTSLLTSSRGHAGATELVALTAAGAHVTVA 1517
 QY 61 CDTNSPQLOQLANTIPROHPLTVITHTAGILDDATITNTPTQANVLAQAKHSAHLH 120
 DB 1518 CDTTDPQLATLTLTTPTEHPLTVITHTAGILDDATITNTPTQANVLAQAKHSAHLH 1577
 QY 121 QLTQHTPLTFAVLVSSAAATFGAGQANYAAAYDALAHHRHTHLPTSIAWGTM 178
 DB 1578 HLTLPHTVTFVLVSSAAGQAGTGAQANYAAAYDALAHHRHTHLPTSIAWGTM 1635

RESULT 5

Q9ADL6 PRELIMINARY; PRT; 6315 AA.

AC Q9ADL6 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Soraphen polyketide synthase A.
 DB SOR.
 OS Polyangium cellulosum.
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
 OC Myxococcales; Sorangineae; Polyangiales; Polyangium.
 OC NCBI_TaxID=56;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SO CE26;
 RA Ligon J.M., Hill S., Beck J., Zirkle R., Molnar I., Zawodny J.,
 RA Money S., Schupp T.;
 RT "Characterization of the biosynthetic gene cluster for the antifungal
 RT polyketide soraphen A from Sorangium cellulosum So CE26.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U24241; AKI19883.1; -.
 DR HSP; P25715; IMLA.
 DR InterPro: IPR001227; AC transferase.
 DR InterPro: IPR001297; ADH_short.
 DR InterPro: IPR002085; adh_zn_family.
 DR InterPro: IPR004410; PabD.
 DR InterPro: IPR00794; ketoacyl-synt.
 DR InterPro: IPR003880; Ppantne_attach.
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00698; Acyl_transferase; 4.
 DR Pfam: PF00106; adh_short; 3.
 DR Pfam: PF00107; adh_zinc; 2.
 DR Pfam: PF00109; ketoacyl-synt; 3.
 DR Pfam: PF02801; ketoacyl-synt_C; 3.
 DR Pfam: PF00550; pp-binding; 4.
 DR TIGRfam: TIGR00128; fadD; 4.
 DR PROSITE: PS50075; ACP_DOMAIN; 4.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 3.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 3.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_4.
 KW Phosphopantetheine; transferase.

SQ SEQUENCE 6315 AA; 671301 MW; 7B3FB977006C12FE CRC64;

Query Match 52.6%; Score 501; DB 2; Length 6315;
 Best local Similarity 53.4%; Pred. No. 7.8e-36;

Matches 95; Conservative 25; Mismatches 58; Indels 0; Gaps 0;

QY 1 GTTLTGGTGAATLTHLTHLTHQPTQHLLTSRTGPHRPHAOHLTTQLOQKIHLLTTT 60
 DB 3803 GTVLITGGTGLGALLARHLVTRHGVTSLLTSSRGHAGATELVALTAAGAHVTVA 3862
 QY 61 CDTNSPQLOQLANTIPROHPLTVITHTAGILDDATITNTPTQANVLAQAKHSAHLH 120
 DB 3863 CDAADPRALQALLDLSIPSAHPLTAVVAAGALDDGLGAMSPERIDRVFAPKDAAMHLH 3922
 QY 121 QLTQHTPLTFAVLVSSAAATFGAGQANYAAAYDALAHHRHTHLPTSIAWGTM 178
 DB 3923 ELTQDPRLAFLVLFSSAAGVLSFGQSNYAAAYDALAHHRHTHLPTSIAWGTM 3980

RESULT 6

Q9LAX2 PRELIMINARY; PRT; 5435 AA.

AC Q9LAX2 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE NYSJ.
 GN NYSJ.
 OS Streptomyces noursei.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=1971;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 11455;
 RX MEDLINE=20334850; PubMed=10873841;
 RA Brautaset T., Sekurova O.N., Sletta H., Ellingsen T.E., Strom A.R.,
 RA Valla S., Zotchev S.B.;
 RT "Bioynthesis of the polyene antifungal antibiotic nystatin in
 RT Streptomyces noursei ATCC 11455: analysis of the gene cluster and
 RT deduction of the biosynthetic pathway.";
 RL Chem. Biol. 7:395-403(2000).
 DR EMBL: AF263912; AAF7167.1; -.
 DR HSP; P25715; IMLA.
 DR InterPro: IPR001227; AC transferase.
 DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR002085; adh_zn_family.
 DR InterPro: IPR004410; PabD.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR000794; ketoacyl-synt.
 DR InterPro: IPR003880; Ppantne_attach.
 DR Pfam: PF00698; Acyl_transferase; 3.
 DR Pfam: PF00106; adh_short; 1.
 DR Pfam: PF00107; adh_zinc; 1.
 DR Pfam: PF00109; ketoacyl-synt; 3.
 DR Pfam: PF02801; ketoacyl-synt_C; 3.
 DR Pfam: PF00550; pp-binding; 3.
 DR TIGRfam: TIGR00128; fadD; 3.
 DR PROSITE: PS50075; ACP_DOMAIN; 3.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 3.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 3.
 KW Phosphopantetheine; transferase.
 SQ SEQUENCE 5435 AA; 562659 MW; AA55465DF087A38C CRC64;

Query Match 52.5%; Score 500; DB 2; Length 5435;
 Best local Similarity 55.1%; Pred. No. 8e-36;

Matches 98; Conservative 16; Mismatches 64; Indels 0; Gaps 0;

QY 1 GTTLTGGTGAATLTHLTHLTHQPTQHLLTSRTGPHRPHAOHLTTQLOQKIHLLTTT 60
 DB 1740 GTVLITGGTGLGALLARHLVTRHGVTSLLTSSRGHAGATELVALTAAGAHVTVA 1799

QY 61 CDTNSPDQLOQLNTIPQHPHTVHTAGILDDATLTNTLPQTQNNVLRKAKASHLLH 120
 DB 1800 CDVADRFALALATLTPAEHPLTVAVHTAGVLDGDTLTALNPDALATVLRKADAAWHLL 1859
 QY 121 QLTQHTPLTAFLVLYSSAAATGAGGQANYAAYLDALAHNRHTHLLPATSTANGTW 178
 DB 1860 DLTRHLDLAFVLYSSAGVGGGQANYAAGNTFLDALAHRAHGLPATSTLANGW 1917

RESULT 7

Q914W3 PRELIMINARY; PRT; 11096 AA.
 AC Q914W3;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE MYC.
 GN MYC.
 OS Streptomyces noursei.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyicinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1971;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 11455;
 RX MEDLINE=20334850; PubMed=10873841;
 RA Brautaset T., Sekurova O.N., Sletta H., Ellingsen T.E., Strom A.R.,
 Valla S., Zotchev S.B.,
 RT "Biosynthesis of the polyene antifungal antibiotic nystatin in
 Streptomyces noursei ATCC 11455: analysis of the gene cluster and
 RT deduction of the biosynthetic pathway.",
 RL Chem. Biol. 7:395-403(2000).
 DR EMBL: AF263912; AAF1776.1; -.
 DR HSSP: P25715; IMLA
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR002085; Adh_zn_family.
 DR InterPro: IPR004410; Fabd.
 DR InterPro: IPR000794; ketoacyl-synt.
 DR InterPro: IPR003880; Pantne attach.
 DR Pfam: PF00698; Acyl_transf; 6.
 DR Pfam: PF00107; adh_zinc; 1.
 DR Pfam: PF00109; ketoacyl-synt; 6.
 DR Pfam: PF02801; ketoacyl-synt C; 6.
 DR Pfam: PF00550; pp-binding; 6.
 DR TIGRFAMs: TIGR00128; fabd; 3.
 DR PROSITE: PS50075; ACP DOMAIN; 6.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 6.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 5.
 KW Phosphopantetheine; transferase.
 SQ SEQUENCE 11096 AA; 1150415 MW; 776CAEAFCAE51DD CRC64;

Query Match 50.7%; Score 483; DB 2; Length 11096;
 Best Local Similarity 54.5%; Pred. No. 6,4e-34;
 Matches 97; Conservative 19; Mismatches 62; Indels 0; Gaps 0;

QY 1 GTTLITGGTGAATLTHLTHHTTQPTQHLILTSRTGPHRPAQHLTQLOQKIHLLTTT 60
 DB 1381 GTVLTIGTGGLGAVLRAHIVATGVRHLLASRGRAADGADLTAEELGAGTVIAA 1440
 QY 61 CDTNSPDQLOQLNTIPQHPHTVHTAGILDDATLTNTLPQTQNNVLRKAKASHLLH 120
 DB 1441 CDVADRFALALATLTPAEHPLTVAVHTAGVLDGDTLTALNPDALATVLRKADAAWHLL 1500
 QY 121 QLTQHTPLTAFLVLYSSAAATGAGGQANYAAYLDALAHNRHTHLLPATSTANGTW 178
 DB 1501 DLTRHLDLAFVLYSSAGVGGGQANYAAGNTFLDALAHRAHGLPATSTLANGW 1558

RESULT 8
 Q93JNX8 PRELIMINARY; PRT; 5644 AA.
 AC Q93JNX8;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE AmphJ.
 GN AMPHJ.
 OS Streptomyces nodosus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyicinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=40318;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Caffrey P., Lynch S.V., Flood E.M., Finnan S.M., O'Leary M.;
 RT "The amphotericin biosynthetic gene cluster from Streptomyces
 RT nodosus.",
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF57202; AAK73502.1; -.
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR002085; Adh_zn_family.
 DR InterPro: IPR004410; Fabd.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR000794; ketoacyl-synt.
 DR InterPro: IPR003880; Pantne attach.
 DR Pfam: PF00698; Acyl_transf; 3.
 DR Pfam: PF00107; adh_zinc; 1.
 DR Pfam: PF00109; ketoacyl-synt; 3.
 DR Pfam: PF02801; ketoacyl-synt C; 3.
 DR Pfam: PF00550; pp-binding; 3.
 DR TIGRFAMs: TIGR00128; fabd; 3.
 DR PROSITE: PS50075; ACP DOMAIN; 3.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_3.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_3.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_3.
 KW Phosphopantetheine.
 SQ SEQUENCE 5644 AA; 585775 MW; CF6A35E3BA0D2658 CRC64;

Query Match 50.5%; Score 481; DB 2; Length 5644;
 Best Local Similarity 51.7%; Pred. No. 4,3e-34;
 Matches 92; Conservative 24; Mismatches 62; Indels 0; Gaps 0;

QY 1 GTTLITGGTGAATLTHLTHHTTQPTQHLILTSRTGPHRPAQHLTQLOQKIHLLTTT 60
 DB 1420 GTVLTIGTGGLASALARRHVAEHDIRHLLTSRRGPDAGADLVQALAELEAEARVAA 1479
 QY 61 CDTNSPDQLOQLNTIPQHPHTVHTAGILDDATLTNTLPQTQNNVLRKAKASHLLH 120
 DB 1480 CDVADRFALALATLTPAEHPLTVAVHTAGVLDGDTLTALNPDALATVLRKADAAWHLL 1539
 QY 121 QLTQHTPLTAFLVLYSSAAATGAGGQANYAAYLDALAHNRHTHLLPATSTANGTW 178
 DB 1540 DLTRHLDLAFVLYSSAGVGGGQANYAAGNTFLDALAHRAHGLPATSTLANGW 1597

RESULT 9

O30768 PRELIMINARY; PRT; 1879 AA.
 ID O30768;
 AC O30768;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Polypeptide synthase module 7.
 GN NID5.
 OS Streptomyces caelestis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyicinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=16816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL-2821;
 RX MEDLINE=98053867; PubMed=9393718;
 RA Kakavas S.J., Katz L., Stasael D.;
 RT "Identification and characterization of the nidamycin polyketide
 RT synthase genes from Streptomyces caelestis.",
 RL J. Bacteriol. 179:7515-7522(1997).
 DR EMBL: AF016585; AAC46028.1; -.
 SQ SEQUENCE 1879 AA; 1950415 MW; 776CAEAFCAE51DD CRC64;

ID	O9EWAL		PRELIMINARY;	PRT;	9507 AA.
AC	O9EWAL:				
DT	01-MAR-2001	(TREMBLrel. 16,	Created)		
DT	01-MAR-2001	(TREMBLrel. 16,	Last sequence update)		
DT	01-JUN-2002	(TREMBLrel. 21,	Last annotation update)		
DE	pImS2 protein.				
GN	PIMS2				
OS	Streptomyces natalensis.				
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;				
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.				
OX	NCBI_TaxID=68242;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20547809; PubMed=11094342;				
RA	Aparicio J.F., Foces R., Mendes M.V., Olivera N., Martin J.F.;				
RT	"A complex multienzyme system encoded by five polyketide synthase				
RT	genes is involved in the biotransformation of the 26-membered polyene				
RU	macrolide pimaricin in Streptomyces natalensis.";				
RU	Chem. Biol. 7:895-905(2000).				
DR	EMBL, AJ278573; CAC20921.1; --				
DR	HSSP; P25715; IMLA.				
DR	InterPro; IPR001227; Ac transferase.				
DR	InterPro; IPR002198; ADH short.				
DR	InterPro; IPR004410; FADd.				
DR	InterPro; IPR001899; Gram_pos_anchor.				
DR	InterPro; IPR000794; Ketoacyl-synt.				
DR	InterPro; IPR001005; MyD_DNA_binding.				
DR	InterPro; IPR003880; Pantane_attach.				
DR	Pfam; PF00698; Acyl_transf. 6.				
DR	Pfam; PF00106; adh_short; 3.				
DR	Pfam; PF00109; ketoacyl-synt.; 6.				
DR	Pfam; PF02801; ketoacyl-synt-C; 6.				
DR	Pfam; PF005501; pp-binding; 6.				
DR	TIGRFAMS; TIGR00128; fadD; 6.				
DR	PROSITE; PS50075; ACID DOMAIN; 6.				
DR	PROSITE; PS00343; B_KETOACYL_SYNTHASE; 6.				
DR	PROSITE; PS00037; MYB_1; UNKNOWN_1.				
DR	PROSITE; PS00012; PHOSPHOPANTETHEINE; 5.				
KW	Phosphopantetheine.				
SQ	SEQUENCE 9507 AA; 994228 MW; 57BD80C6AD37E89F CRC64;				
Query Match 49.5%; Score 471.5; DB 2; Length 9507;					
Best Local Similarity 51.7%; Pred. No. 5.8e-33;					
Matches 92; Conservative 21; MisMatches 64; Indels 1; Gaps 1;					
OY	1 GTTLTGTGATATLTHNLTTHTOPTHLLTSRGPTTPAHQHTLTLOOKGHILTTT 60				
DB	6025 GTVLTGGTGAGLGGVAAFWL-AEAGAETHLLVSRRGTDAFGAGEIAAIIEAGARVTVA 60833				
OY	61 CDTSNPDLQOLQNLNTIPPOHPPLTVIHTAGILDADATLTNLITPTQLNVIRAKASAHILL 120				
DB	6084 CDTARDRLAANVALAIIPDEHPLTAFTFRNGIVTDOSTLTITPECPASTVRKRYATATRH 61433				
OY	121 QLTQHPTLATAVLYSSAAATGCAPGOANYAANAVALDALAHHRHTHLPATSIANGTW 178				
DB	6144 ETRREODLSAEFLFESSVAGTIGAFPGQNYAAGNAFLDAFAARRAHGLPATSIAMGPW 6201				
RESULT 13					
ID	O93NX7		PRELIMINARY;	PRT;	2034 AA.
AC	O93NX7:				
DT	01-DEC-2001	(TREMBLrel. 19,	Created)		
DT	01-DEC-2001	(TREMBLrel. 19,	Last sequence update)		
DT	01-JUN-2002	(TREMBLrel. 21,	Last annotation update)		
DE	Amphk.				
GN	AMPBK.				
OS	Streptomyces nodosus.				
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;				
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.				
OX	NCBI_TaxID=40318;				

```

RN      [1]
RP      SEQUENCE FROM N.A.
RA      Caffrey P., Lynch S.V., Flood E.M., Finnian S.M., O'Leary M.;
RT      "The amphitericin biosynthetic gene cluster from Streptomyces
RL      nodosus.";
RL      Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF357202; AAK73503.1; -
DR      InterPro; IPR001227; Ac transferase.
DR      InterPro; IPR002198; ADH_short.
DR      InterPro; IPR000794; ketoacyl-synt.
DR      InterPro; IPR003880; Pantine atch.
DR      InterPro; IPR003379; Ser esterase.
DR      InterPro; IPR001031; Thioesterase.
DR      Pfam; PF00698; Acyl_transf. 1.
DR      Pfam; PF00106; adh_short; 1.
DR      Pfam; PF00109; ketoacyl-synt. 1.
DR      Pfam; PF02801; ketoacyl-synt. C; 1.
DR      Pfam; PF00550; pp-binding; 1.
DR      Pfam; PF00975; Thioesterase; 1.
DR      PROSITE; PS50075; ACP_DOMAIN; 1.
DR      PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
KW      phosphopantetheine.
SQ      SEQUENCE 2034 AA; 212063 MW; 83E962B7DC3D5747 CRC64;

Query Match          49.4%; Score 471; DB 2; Length 2034;
Best Local Similarity 52.2%; Pred. No. 1e-33;
Matches 93; Conservative 19; Mismatches 66; Indels 0; Gaps 0;

QY      1 GTTLITGGTATLATHLTHLTHHPTHOPTOHLILTSRTGPHTHAOHLTTLOOOGIHLTTT 60
DB      1386 GTVLITGGTGLGAAALRAHLVTSGARHLLASRGFDARETSLVAELTGLGQIVYSA 1445
QY      61 CDTSNPDOLOLLNTTTPRQHELTTVYHTAGILDDATLTNTLPTQANNVRAKAHSAHLH 120
DB      1446 CDVGRDRAVDPTLVAASVPAEHPDLTVAVHTAGVLDALTGSLTPEQIAGVLPKADALHLH 1505
QY      121 QLTQHTPLTAAVLYSSAAATGAGCGAANYAANAAYDALAHNRHTHLPATSTINGTW 178
DB      1506 EATLGODLAARVLYSSISGVIGGQANYAANAASLDALHRRRAAGLPGLSIAMGPW 1563

RESULT 14
Q9EX53 PRELIMINARY; PRT; 3576 AA.
AC      Q9EX53;
DT      01-MAR-2001 (TREMblrel. 16, Created)
DT      01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT      01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE      Putative type I polyketide synthase.
SC      SC06274 OR SC2C4.04C.
OS      Streptomyces coelicolor.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX      NCBI_TaxID=1902;
[1]
RN      SEQUENCE FROM N.A.
RC      STRAIN=A3(2);
RA      Seeger K.J., Harris D.;
RL      Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN      SEQUENCE FROM N.A.
RC      STRAIN=A3(2);
RA      Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL      Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
[3]
RN      SEQUENCE FROM N.A.
RC      STRAIN=A3(2);
RA      MEDLINE=97000351; PubMed=8843436;
RA      Redenbach M., Kleiser H.M., Denapate D., Eichner A., Cullum J.,
RA      Kinashi H., Hopwood D.A.;
RT      "A set of ordered cosmids and a detailed genetic and physical map for
RT      the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL      Mol. Microbiol. 21:77-96(1996).

```

```

RT      avermitilis: Deducing the ability of producing secondary
KL      Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR      EMBL; AB070940; BAB69193.1; -.
DR      InterPro; IPR001227; Ac transferase.
DR      InterPro; IPR002198; ADH short.
DR      InterPro; IPR004410; Fabb.
DR      InterPro; IPR001899; Gram_pos_anchor.
DR      InterPro; IPR000794; Ketoacyl-synt.
DR      InterPro; IPR003880; Pantne_attach.
DR      Pfam; PF00698; Acyl_transf; 3.
DR      Pfam; PF00106; adh_short; 1.
DR      Pfam; PF00109; ketoacyl-synt; 3.
DR      Pfam; PF02801; ketoacyl-synt_C; 3.
DR      Pfam; PF00550; dp-binding; 3.
DR      TIGRFAMs; TIGR00128; fabd; 3.
DR      PROSITE; PS50075; ACP_DOMAIN; 3.
DR      PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_3.
DR      PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR      PROSITE; PS00012; PHOSPHOPANTHETINE; UNKNOWN_2.
KW      Phosphopantetheine.
SQ      SEQUENCE 4840 AA; 504325 MW; 169BDA3589AD284F CRC64;

Query Match      49.4%; Score 471; DB 2; Length 4840;
Best Local Similarity 50.6%; Pred. NO.2.9e-33;
Matches 90; Conservative 23; Mismatches 65; Indels 0; Gaps 0;

OY      1 GTTLITGTCGATLATHLTTHLTTTHQPTOHLLTSRTGPTPHAOHLTTTLOOKGIIHTTT 60
DB      4401 GTVVATGATGCGTGTETIRHATAGVGNHLLLSRSGPDDPRAAELGKVLAEAGDATTFA 4460
OY      61 CDTSPNDLOOLNTLTIPOHPLTTVIHTAGILDDATLTNLPTQANNVLRKASHAHLH 120
DB      4461 CDTADRAALADALADVPAREHPVAVVHIAGVDDGVTLLTPRIDTVLRPKAAQHLH 4520
OY      121 QLTQHTPLTAQVLYSSAAATFGAGQANVYAAANVYLDALAHHRHTHLPATSTIAGTW 178
DB      4521 ELTADLGLSHFVLPSSGVGTFGGAGQANVYAAANVFLDALHRRRAADLPATSIAMGLM 4578

Search completed: June 17, 2003, 13:07:54
Job time : 9.22676 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:52:22 ; Search time 1.94115 Seconds

(without alignments)
3803.313 Million cell updates/sec

Title: US-09-914-286-3_COPY_1715_1892

Perfect score: 953
Sequence: 1 GTTLITGTCGALATHLTHL.....LAHHRHTHLPAISIAWGTW 178

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database: SwIsProt_40.*
Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	391	41.0	3567	1	ERY2_SACER
2	389.5	40.9	3491	1	ERY1_SACER
3	326	34.2	3172	1	Q03133 saccharopol
4	295	31.0	3519	1	OL56_STRAT
5	271.5	28.5	2110	1	MCAS_MYCBO
6	217	22.8	1774	1	MSAS_PENPA
7	217	22.8	1876	1	PPSA_MYCTU
8	203	21.3	4447	1	PKSK_BACSU
9	201.5	21.1	2505	1	PAS_RAT
10	196.5	20.6	1538	1	PPSB_MYCTU
11	195	20.5	4427	1	PKSL_BACSU
12	191.5	20.1	2511	1	PAS_CHICK
13	182.5	19.2	2504	1	PAS_HUMAN
14	179.5	18.8	838	1	PAS_MOUSE
15	146	15.3	4273	1	PKSM_BACSU
16	111	11.6	247	1	YGPF_ECOLI
17	109.5	11.5	255	1	PABG_MYCAT
18	108.5	11.4	247	1	PABG_MYCTU
19	106	11.1	247	1	YD50_MYCTU
20	105.5	11.1	260	1	TRN2_DATST
21	101.5	10.7	247	1	PABG_CHLTR
22	100.5	10.5	268	1	TRNH_DATST
23	99.5	10.4	244	1	PABG_SALT
24	98.5	10.3	244	1	PABG_ECOLI
25	97	10.2	246	1	PHBB_CHRVI
26	96.5	10.1	246	1	PABG_BACSU
27	96	10.1	606	1	HMID_DROAN
28	94.5	9.9	255	1	PABG_MYCSM
29	93.5	9.8	260	1	TRN2_HYONI
30	92.5	9.7	244	1	PABG_VIBCH
31	91.5	9.6	241	1	PABG_RIPER
32	91.5	9.6	245	1	NODG_RHIS3
33	90.5	9.5	334	1	GALE_BACHD

34	90	9.4	372	1	3BH3_RAT	P27364 ratius norv
35	89.5	9.4	245	1	NODG_RHIME	P06234 rhizobium m
36	89.5	9.4	246	1	NODG_AZOBR	P17611 azobryllu
37	89.5	9.4	252	1	YCIR_ECOLI	P31808 escherichia
38	89	9.3	244	1	PABG_BUCAT	P57432 buchnera ap
39	89	9.3	609	1	OPA_DROME	P39768 drosophila
40	89	9.3	1301	1	PTP9_DROME	P35832 drosophila
41	88	9.2	247	1	PABG_PSEAE	O54438 pseudomonas
42	88	9.2	906	1	FOX2_CANTR	P22414 candida tro
43	87.5	9.2	242	1	PABG_HAEIN	P43713 haemophilus
44	87.5	9.2	248	1	PABG_CHIMU	O9PK77 chlamydia m
45	87.5	9.2	1829	1	Y296_HUMAN	O15015 homo sapien

ALIGNMENTS

RESULT 1
ID ERY2_SACER STANDARD; PRT; 3567 AA.
AC Q03132; O54096;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Erythronolide synthase, modules 3 and 4 (EC 2.3.1.94) (ORF 2) (6-
DE deoxyerythronolide B synthase II) (DEBS 2).
GN ERYA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Actinobacteria; Actinobacteridia (class); Actinobacteridae;
OC Actinomycetales; Pseudonocardiales; Pseudonocardaceae;
OC Saccharopolyspora.
OX NCBI_TaxId=1836;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9122065; PubMed=2024119;
RA Donadio S., Steaver M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT "Modular organization of genes required for complex polyketide
RT biosynthesis.";
RL Science 252:675-679 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=NRRL 2338;
RA MEDLINE=92155230; PubMed=1740151;
RA Bevil D.J., Cortes J., Haydock S.F., Leadlay P.F.;
RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea.
RT Cloning of the structural gene, sequence analysis and inferred domain
RT structure of the multifunctional enzyme.";
RL Eur. J. Biochem. 204:39-49 (1992).
CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
CC deoxyerythronolide B.
CC -1- COPFACTOR: NADP, CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHEINES.
CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
CC BIOSYNTHESIS.
CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH
CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3
CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
CC SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS
CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
CC RESPECTIVELY.
CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),
CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS) AND ACYL CARRIER
CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
CC OF THE FULL-LENGTH CHAIN.
CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.

CC
 DR EMBL, M63677; AAA26494.1; -;
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR002085; Adh zn family.
 DR InterPro: IPR000794; Ketocacyl-synt.
 DR InterPro: IPR003880; Pplantn_attach.
 DR Pfam: PF00107; adh_zinc; 1.
 DR Pfam: PF00109; ketocacyl-synt; 2.
 DR Pfam: PF00550; pp-binding; 2.
 DR Pfam: PF00698; Acyl transfer; 2.
 DR Pfam: PF02801; ketocacyl-synt C; 2.
 DR PROSITE, PS00012; PHOSPHOPANTHETHEINE; 2.
 DR PROSITE, PS00606; B KETOACYL SYNTHASE; 2.
 DR TRANSFERASE; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KM Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1484
 FT DOMAIN 1485 3567
 FT DOMAIN 27 488
 FT DOMAIN 559 884
 FT DOMAIN 1130 1301
 FT DOMAIN 1397 1467
 FT DOMAIN 1485 1943
 FT DOMAIN 2013 2336
 FT DOMAIN 2383 3066
 FT DOMAIN 3139 3322
 FT DOMAIN 3415 3485
 FT ACT_SITE 202 202
 FT ACT_SITE 651 651
 FT BINDING 1430 1430
 FT ACT_SITE 1661 1661
 FT ACT_SITE 2115 2115
 FT NP_BIND 2961 2978
 FT NP_BIND 3142 3157
 FT BINDING 3448 3448
 FT CONFLICT 438 438
 FT CONFLICT 480 480
 FT CONFLICT 1241 1241
 FT CONFLICT 2664 2664
 SQ SEQUENCE 3567 AA; 374413 MW; E86284F4738A0C0 CRC64;

Query Match 41.0%; Score 391; DB 1; Length 3567;
 Best Local Similarity 46.1%; Pred. No. 5.6e-26;
 Matches 82; Conservative 24; Mismatches 70; Indels 2; Gaps 1;

QY 1 GTTLITGCGALATHLTHHLLTTHQPRQHLITSRGHTPHAOHLTTOLQKGIHLITTT 60
 DB 3141 GTVLTIGTGTLGRLRLHLYTERGVHLLVSRGADAPQSDRLRAIEDLGASATIAA 3200
 QY 61 CDTSNPQLOOLANTTIPPOHPLTVITAGILDDATLTNTLPTQINNVNRAKASAHLLH 120
 DB 3201 CDTADRALALADLGLP--RLPTGVVAAGVLDGLVTSIDEPVEQVLAQVDAAMNHL 3258
 QY 121 QLTQHTPLTAVLYSSAAATFGAGQANVAAAYLDALAHHRHTHLPTATSLAWGTW 178
 DB 3259 ELTANTGLSFFVLFFSSAASVLAGGQGVAAANESLNAALARTRTGLPAKALGMGIM 3316

RESULT 2
 ID ERY1_SACER STANDARD; PRT; 3491 AA.
 AC 003131;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Erythronolide synthase, modules 1 and 2 (EC 2.3.1.94) (ORF 1) (6-
 DE deoxyerythronolide B synthase 1) (DBS 1).
 FT ERYA.

OS Saccharopolyspora erythraea (Streptomyces erythraeus).
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Pseudonocardineae; Pseudonocardaceae;
 CC Saccharopolyspora.
 OK NCBI_TaxID=1836;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220065; Pubmed=2024119;
 RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.U., Katz L.;
 RT "Saccharopolyspora erythraea.",
 RL Gene 126:147-151(1993).
 CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
 CC deoxyerythronolide B.
 CC -1- COFACTOR: NADP, CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTHETHEINES.
 CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
 CC BIOSYNTHESIS.
 CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH
 CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3
 CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
 CC SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS
 CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
 CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
 CC RESPECTIVELY.
 CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),
 CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER
 CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
 CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
 CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
 CC OF THE FULL-LENGTH CHAIN.
 CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
 CC -1- SIMILARITY: CONTAINS 3 ACYL CARRIER DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC
 CC EMBL, M63676; AAA26493.2; -;
 CC EMBL, L07626; AAA26504.1; -;
 DR HSSP: P25715; IMLA.
 DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR000794; ketocacyl-synt.
 DR InterPro: IPR003880; Pplantn_attach.
 DR Pfam: PF00106; adh_short; 1.
 DR Pfam: PF00109; ketocacyl-synt; 2.
 DR Pfam: PF00550; pp-binding; 3.
 DR Pfam: PF00698; Acyl transfer; 3.
 DR Pfam: PF02801; ketocacyl-synt C; 2.
 DR PROSITE, PS00012; PHOSPHOPANTHETHEINE; 3.
 DR PROSITE, PS00606; B KETOACYL SYNTHASE; 2.
 DR PROSITE, PS00075; ACP DOMAIN; 3.
 DR TRANSFERASE; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KW Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1972
 FT DOMAIN 1979 3491
 FT DOMAIN 1 375
 FT DOMAIN 414 484
 FT DOMAIN 503 961
 FT DOMAIN 1030 1356
 FT DOMAIN 1611 1794
 FT DOMAIN 1888 1958

Query	1	GTTLTGCTGALATLTHLTTTHOFTQHLTLSTRTGTPRHQHLTLOOKGHLITTT	60
Db	1613	GTVALTGCTGCGVGOIARWL-ARRGAPHLILVSRSGPADGAGELVALLEALGARTTVA	167121
Qy	61	CDTSNPDOLOQLNTIPRQHLLTYTHNAGLDDDTLNTLPQTQANLTKAKSAHLH	120
Db	1672	CDVTDRESRELLGGIGDVPISAFHAAATLDDSTVTLTGERTERSRKVLGARHL	173131
Qy	121	QLQHTPTATAVLVSSAATFGAPQOAVYAANAVLDAHLAHRPHHLPATSIAGTW	178
Db	1732	ELTVELDLTATVLPSSPASAGARBLGGYATFGNAYLDLDAQRRSDGLPATYAVKGTW	1789

Query Match	40.9%	Score 389.5	DB 1	Length 3491
Beat Local Similarity	44.4%	Pred. No. 7.4e-26		
Matches	79	Conservative	25	Mismatches 73, Indels 1, Gaps 1

FT	DOMAIN	1979	2441	ACYL-KETOACYL SYNTHASE 2.
FT	DOMAIN	2507	2854	BETA-TRANSPEASE (AT) 3.
FT	DOMAIN	3055	3237	BETA-KETOACYL REDUCTASE 2.
FT	DOMAIN	3334	3404	ACYL CARRIER (ACP) 3.
FT	ACT_SITE	145	145	ACYL-ENZYME INTERMEDIATE.
FT	BINDING	447	447	PHOSPHOPANTHETHEINE (BY SIMILARITY).
FT	ACT_SITE	677	677	BETA-KETOACYL SYNTHASE.
FT	ACT_SITE	1128	1128	ACYL-ENZYME INTERMEDIATE.
FT	NP BIND	1614	1660	NADP.
FT	BINDING	1921	1921	PHOSPHOPANTHETHEINE (BY SIMILARITY).
FT	ACT_SITE	2148	2148	BETA-KETOACYL SYNTHASE.
FT	ACT_SITE	2598	2598	ACYL-ENZYME INTERMEDIATE.
FT	NP BIND	3058	3104	NADP.
FT	BINDING	3367	3367	PHOSPHOPANTHETHEINE (BY SIMILARITY).
SO	SEQUENCE	3491 AA	365022 MW	682BFC32C90FAC4 CRC64

CS	Saccharopolyspora erythraea (Streptomyces erythraeus).
OC	Bacteria; Actinobacteria; Actinobacteria (Class); Actinobacteridae;
OC	Actinomycetales; Pseudonocardiales; Pseudonocardiaceae;
OC	Saccharopolyspora.
OX	NCBI_TaxId=1836;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=NRRL 2338;
RC	MEDLINE=92155230; PubMed=2234082;
RA	Bevilitt D.J., Cortes S.F., Haydock S.F., Leadlay P.F.;
RT	"An unusually large multifunctional polypeptide in the erythromycin-
RL	Nature 348:176-178 (1990).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	MEDLINE=91220065; PubMed=2024119;
RA	Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT	"Modular organization of genes required for complex polyketide
RL	Science 252:675-679 (1991).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN=NRRL 2338;
RC	MEDLINE=92155230; PubMed=1740151;
RA	Bevilitt D.J., Cortes S.F., Haydock S.F., Leadlay P.F.;
RT	"6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea.
RL	Cloning of the structural gene, sequence analysis and inferred domain
CC	structure of the multifunctional enzyme.";
CC	Bur. J. Biochem. 204:39-49 (1992).
CC	-1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-

CC	deoxyerythronolide B.
CC	-1- COFACTOR: NADP. CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.
CC	-1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
CC	-1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
CC	-1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR), DEHYDATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.
CC	-1- SIMILARITY: TO FATTY ACID SYNTHASE (PAS).
CC	-1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk).
CC	-----
DR	EMBL; X56107; CAA39583.1; -
DR	EMBL; M63677; AAA26495.1; -
DR	EMBL; X62569; CAA44449.1; -
DR	HSSP; P00101; ICCH.
DR	InterPro; IPR002198; ADH_short.
DR	InterPro; IPR001227; Ac transferase.
DR	InterPro; IPR000794; Ketoacyl-synt.
DR	InterPro; IPR003880; Pantane_attach.
DR	InterPro; IPR001031; Thioesterase.
DR	Pfam; PF00106; adh_short; 1.
DR	Pfam; PF00109; ketoacyl-synt; 2.
DR	Pfam; PF00550; pp-binding; 2.
DR	Pfam; PF00698; Acyl transfer; 2.
DR	Pfam; PF00975; Thioesterase; 1.
DR	Pfam; PF02801; ketoacyl-synt_C; 2.
DR	PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
DR	PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
DR	PROSITE; PS0075; ACP DOMAIN; 2.
KW	Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat; Phosphopantetheine; Multifunctional enzyme.
KW	Phosphopantetheine; Multifunctional enzyme.
FT	DOMAIN 1 1484
FT	DOMAIN 1485 3172
FT	DOMAIN 3172 484
FT	DOMAIN 484 878
FT	DOMAIN 878 1298
FT	DOMAIN 1298 1394
FT	DOMAIN 1394 1464
FT	DOMAIN 1464 1954
FT	DOMAIN 1488 1954
FT	DOMAIN 2021 2335
FT	DOMAIN 2255 2735
FT	DOMAIN 2821 2891
FT	DOMAIN 2926 3172
FT	ACT_SITE 199 199
FT	ACT_SITE 643 643
FT	NP_BIND 1118 1164
FT	BINDING 1427 1427
FT	ACT_SITE 1661 1661
FT	ACT_SITE 2112 2112
FT	NP_BIND 2557 2605
FT	BINDING 2854 2854
FT	CONFLICT 231 231
FT	CONFLICT 240 240
FT	CONFLICT 289 289
FT	CONFLICT 493 493
FT	CONFLICT 517 517
FT	CONFLICT 493 517

FT CONFLICT 510 510 A -> R (IN REF. 2).
 FT CONFLICT 513 513 M -> W (IN REF. 2).
 FT CONFLICT 525 525 E -> D (IN REF. 2).
 FT CONFLICT 536 536 R -> G (IN REF. 2).
 FT CONFLICT 547 551 GPNRP -> ARTR (IN REF. 2).
 FT CONFLICT 553 553 R -> G (IN REF. 2).
 FT CONFLICT 673 673 R -> A (IN REF. 2).
 FT CONFLICT 716 716 MISSING (IN REF. 2).
 FT CONFLICT 734 736 AHK -> GIT (IN REF. 2).
 FT CONFLICT 896 896 R -> ROR (IN REF. 2).
 FT CONFLICT 896 896 R -> RELPYEPOROR (IN REF. 1).
 FT CONFLICT 988 994 GVAAPR -> VLSLSD (IN REF. 2).
 FT CONFLICT 1108 1116 RTHLEPLA -> ARTWSR (IN REF. 2).
 FT CONFLICT 1132 1132 MISSING (IN REF. 1).
 FT CONFLICT 1132 1132 L -> V (IN REF. 2).
 FT CONFLICT 1192 1192 A -> R (IN REF. 2).
 FT CONFLICT 1194 1194 MISSING (IN REF. 2).
 FT CONFLICT 1277 1278 AA -> RR (IN REF. 2).
 FT CONFLICT 1385 1390 LCGORE -> STARR (IN REF. 2).
 FT CONFLICT 1485 1485 MISSING (IN REF. 2).
 FT CONFLICT 1518 1518 G -> R (IN REF. 2).
 FT CONFLICT 1601 1601 V -> L (IN REF. 2).
 FT CONFLICT 1724 1725 LP -> PA (IN REF. 2).
 FT CONFLICT 1732 1732 Q -> L (IN REF. 2).
 FT CONFLICT 1732 1732 GPAEG -> ARRA (IN REF. 2).
 FT CONFLICT 1762 1762 T -> S (IN REF. 2).
 FT CONFLICT 2252 2252 D -> DGAD (IN REF. 2).
 FT CONFLICT 2275 2277 OSP -> AVA (IN REF. 2).
 FT CONFLICT 2408 2408 G -> GR (IN REF. 2).
 FT CONFLICT 2420 2421 LA -> S (IN REF. 2).
 FT CONFLICT 2443 2444 NA -> TH (IN REF. 2).
 FT CONFLICT 2596 2596 A -> G (IN REF. 2).
 FT CONFLICT 2609 2609 P -> A (IN REF. 2).
 FT CONFLICT 2715 2722 RRAEGRA -> ANKVR (IN REF. 1).
 FT CONFLICT 2754 2754 D -> E (IN REF. 2).
 SQ SEQUENCE 3172 AA; 331474 MW; DBBD5094E77DD5F CRC64;

Query Match 34.2%; Score 326; DB 1; Length 3172;
 Best Local Similarity 42.5%; Pred. No. 2.4e-20;
 Matches 76; Conservative 23; Mismatches 74; Indels 6; Gaps 4;

QY 1 GTTLITGSGALATLTHLTHPTQHLLTSRTGPHPHAOHLTTLOQKGIHLTTT 60
 DB 2557 GTALVTGSGALGSHVARNH-ARCGVEDLVLSRRGVADAPAAELAEALVAGPKTTITA 2615
 QY 61 CDTSNPDLOQLANTIPPO-HEPLTVHTAGILDATLTNLTPTQANNVLRARSHL 119
 DB 2616 CDVADREQLSKLEELRGQRPVTVHTAGVPSRPLHEI--GELSSCAAKVTGARLL 2673
 QY 120 HOITQHTPLTAFVLYSSAATFGAPQOANYAANAVYDALAHHRHTHLPTSIAWGTV 178
 DB 2674 DELCPDA-ETPVLFSSGAGVWGSANLGAISANNAVYDALAHRRARAGRAITSVAMGAW 2730

RESULT 4
 OL56 STRAT STANDARD; PRT; 3519 AA.
 ID OL56 STRAT STANDARD; PRT; 3519 AA.
 AC Q07017;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Oleandomycin polyketide synthase, modules 5 and 6.
 GN ORP.
 OS Streptomyces antibioticus.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1890;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150470; Pubmed=8107683;
 RA Swan D.G., Rodriguez A.M., Valches C., Mendez C., Salas J.A.;
 RT "Characterisation of a Streptomyces antibioticus gene encoding a type
 I polyketide synthase which has an unusual coding sequence.";

RL Mol. Gen. Genet. 242:358-362(1994).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE BIOSYNTHESIS OF THE OLEANDOMYCIN
 CC LACTONE RING.
 CC -1- COPACITOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L09654; AAA19695.1; -
 DR HSSP; P25715; 1MLA.
 DR InterPro; IPR001227; Acyltransferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantane attach.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00109; ketoacyl-synt. 2.
 DR Pfam; PF00550; pp-binding. 2.
 DR Pfam; PF00698; Acyl_transf. 2.
 DR Pfam; PF00975; Thioesterase; 1.
 DR Pfam; PF02801; ketoacyl-synt. C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
 DR PROSITE; PS0075; ACP DOMAIN; 2.
 KW transferase; Acyltransferase; Antibiotic biosynthesis; NADP;
 KW Phosphopantetheine; Multifunctional enzyme; Repeat.
 FT DOMAIN 1 3519 ?
 FT DOMAIN 32 501
 FT DOMAIN 569 890
 FT DOMAIN 1200 1382
 FT DOMAIN 1487 1561
 FT DOMAIN 1686 2156
 FT DOMAIN 2220 2541
 FT DOMAIN 2856 3038
 FT DOMAIN 3141 3215
 FT ACT_SITE ?
 FT BINDING ?
 FT DOMAIN 3270 3519
 FT ACT_SITE 210 210
 FT ACT_SITE 660 660
 FT NP_BIND 1203 1249
 FT BINDING 1524 1524
 FT ACT_SITE 1859 1859
 FT ACT_SITE 2311 2311
 FT NP_BIND 2859 2905
 FT BINDING 3178 3178
 SQ SEQUENCE 3519 AA; 368561 MW; 41A878AAABE81F86 CRC64;

Query Match 31.0%; Score 295; DB 1; Length 3519;
 Best Local Similarity 39.7%; Pred. No. 1.4e-17;
 Matches 71; Conservative 26; Mismatches 78; Indels 4; Gaps 3;

QY 1 GTTLITGSGALATLTHLTHPTQHLLTSRTGPHPHAOHLTTLOQKGIHLTTT 59
 DB 1202 GTALVTGSGALGSHVARNH-VGGADHVLVSRGSGAPGADLVRELGIGARVSR 1260
 QY 60 TCDTSNPDLOQLANTIPPO-HEPLTVHTAGILDATLTNLTPTQANNVLRARSHL 119
 DB 1261 ACVADRVALLRLSLD--GEPTATFHAAGVQSPPLAEISVGEADVAQAQVNVN 1318
 QY 120 HOITQHTPLTAFVLYSSAATFGAPQOANYAANAVYDALAHHRHTHLPTSIAWGTV 178
 DB 1319 GELVDPGCEAFVLFSSNAGVWGSQAVYAANAFLDALAVRRGVGLPATSVAMGAW 1377

RESULT 5
 MCAS MYCBO STANDARD; PRT; 2110 AA.
 ID MCAS MYCBO STANDARD; PRT; 2110 AA.

AC 002251, 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Mycoseroic acid synthase.
 GN MAS.
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxId=1765;
 [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=92406887; PubMed=1527058;
 RA Mathur M., Kolatukudy P.E.;
 RT Molecular cloning and sequencing of the gene for mycoseroic acid
 RT synthase, a novel fatty acid elongating multifunctional enzyme, from
 RT Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guérin.";
 RL J. Biol. Chem. 267:19388-19395(1992).
 CC -1- FUNCTION: CATALYZES THE ELONGATION OF N-FATTY ACYL-COA WITH
 CC METHYLMALONYL-COA (NOT MALONYL-COA) AS THE ELONGATING AGENT TO
 CC FORM MYCOSEROYL LIPIDS.
 CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETHEINE.
 CC -1- SUBUNIT: HOMODIMER WHOSE MONOMERS PROBABLY HAVE A HEAD TO TAIL
 CC ARRANGEMENT.
 CC -1- SUBCELLULAR LOCATION: Membrane-associated.
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M55808; AAA25369.1; -.
 DR PIR: B44110; B44110.
 DR HSSP: P73283; 1ESM.
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR002085; Adh_zn_family.
 DR InterPro: IPR000794; Ketocacyl-synt.
 DR InterPro: IPR003880; Ppantne_attach.
 DR Pfam: PF00107; adh_zinc.1.
 DR Pfam: PF00109; ketocacyl-synt.1.
 DR Pfam: PF00550; pp-binding.1.
 DR Pfam: PF00698; Acyl_transf.1.
 DR Pfam: PF02801; ketocacyl-synt_C.1.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 1.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.
 DR Kary acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 KM Transferase; Hydrolase; Oxidoreductase; Ligase; NADP; Membrane.
 FT DOMAIN 1 430
 FT DOMAIN 533 852
 FT DOMAIN ? ?
 FT DOMAIN ? ?
 FT DOMAIN ? ?
 FT DOMAIN 2026 2096
 FT ACT_SITE 177 177
 FT ACT_SITE 623 623
 FT NP_BIND 1561 1578
 FT NP_BIND 1765 1780
 FT BINDING 2059 2059
 SO SEQUENCE 2110 AA; 225577 MW; CC658215D7155300 CRC64;
 Query Match 28.5%; Score 271.5; DB 1; Length 2110;
 Best Local Similarity 34.6%; Pred. No. 8.7e-16;
 Matches 62; Conservative 33; Mismatches 81; Indels 3; Gaps 3;
 QY 1 GTTLINGGALATHTLTHLTHQPHQLLTSTGPHPHNAHLTLTQKGIHLITTT 60
 DB 1764 GSITITGGGGLGLPFRASKLA--AGCGRIYILRSQPN--PKAOTTEGLAAGADIVVEC 1821
 QY 61 CDTSPDLOQLQTLNTIPPH--PLTVIHTAGILDATLTNLPTQLNNVLRKASHAHL 119

DB 1822 GNAEBPDTRIVSATATGLPLRGVLSAAVEDATLTNITDELIDRMSPFVFGSWNL 1881
 QY 120 HOLCHTPTLAFVLSAATFGAPGQANYAANALDALHHRHTHLPATSIANGTW 178
 DB 1882 HRTTLQQLDWFCFLPSSGALLGSPGQAYVAANSWVDPAHPPRAGCLPVSALANGAW 1940
 RESULT 6
 MSAS PENPA STANDARD; PRT; 1774 AA.
 ID P22367;
 AC 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 6-methylsalicylic acid synthase (EC 2.3.1.-) (MSAS).
 OS Penicillium patulum (Penicillium griseofulvum).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
 OX NCBI_TaxId=5078;
 [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=DSM 62862;
 RX MEDLINE=91006137; PubMed=2209605;
 RA Beck J., Ripka S., Stegner A., Schiltz E., Schweizer E.;
 RT "The multifunctional 6-methylsalicylic acid synthase gene of
 RT Penicillium patulum. Its gene structure relative to that of other
 RT polyketide synthases.";
 RL Eur. J. Biochem. 192:487-498(1990).
 CC -1- FUNCTION: This multifunctional enzyme is a polyketide synthase.
 CC It catalyzes a total of 11 steps by seven different component
 CC enzymes, in the biosynthesis of the antibiotic patulin.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + 3 malonyl-CoA + NADPH =
 CC 6-methylsalicylic acid + NADP(+) + 3 CoA + 3 CO(2) + H(2)O.
 CC -1- PATHWAY: Patulin biosynthesis.
 CC -1- SUBUNIT: HOMOMULTIMER.
 CC -1- INDUCTION: IN THE LATE LOGARITHMIC GROWTH PHASE.
 CC -1- SIMILARITY: WITH MOST POLYKETIDE SYNTHASES, SPECIALLY WITH RAT
 CC FATTY ACID SYNTHASE, AND WITH OTHER ENZYMES SUCH AS LIPASES AND
 CC THIOLEASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X55776; CAA39235.1; -.
 DR PIR: S13178; S13178.
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR000794; Ketocacyl-synt.
 DR InterPro: IPR003880; Ppantne_attach.
 DR Pfam: PF00109; ketocacyl-synt.1.
 DR Pfam: PF00550; pp-binding.1.
 DR Pfam: PF00698; Acyl_transf.1.
 DR Pfam: PF02801; ketocacyl-synt_C.1.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 1.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.
 DR Kary acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 KM Multifunctional enzyme; Oxidoreductase; Antibiotic biosynthesis;
 FT TRANSFERASE; NADP; Phosphopantetheine.
 FT DOMAIN 186 238
 FT DOMAIN 642 676
 FT DOMAIN 1403 1450
 FT DOMAIN 1700 1769
 FT NP_BIND 1419 1424
 FT NP_BIND 204 204
 FT ACT_SITE 653 653
 FT BINDING 1732 1732
 SO SEQUENCE 1774 AA; 190732 MW; 05ED5DD10863F938 CRC64;

```

Query Match      22.8%; Score 217; DB 1; Length 1774;
Best Local Similarity 30.1%; Pred. No. 4.2e-11;
Matches 58; Conservative 28; Mismatches 89; Indels 18; Gaps 3;

QY 1 GTTLLTCGATLATHLTHLTHOPTHLLTSRTGHTPHAOHLTQO-----50
DB 1412 GTVLTTCGLVGLGLEVADFLVEKCARLLISRRALPRRTWQVSDLDPTIAKIRILE 1471
QY 51 QKGIHLITTCDSNPNQLOQLN-----TTPQHPLTVVHTAGIILDDATLNTLPTQL 105
DB 1472 SRGASVHLPLDTKPAVSEQLTALDRLSLP---SVQSVHAGVLDNELVMQTTDAF 1528
QY 106 NNVLAKASAHLLHQLTQHTPLTAVLYSSAATFGAPCOANYAANAVALAHHRHT 165
DB 1529 NRVLAFLAAGALMLHEVFPKPSVDFFMFSSCGNLVFTQASGSGNAFLDTLATRRAR 1588
QY 166 HHLPAISLAWGTW 178
DB 1589 LGDAVASFQWTSW 1601

RESULT 7
PSSA_MYCTU STANDARD; PRT; 1876 AA.
AC Q10977;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DN Phenolphthalein synthase polyketide synthase ppsa.
GN PSSA OR RV2931 OR MT3000 OR MTCY338.20.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
  Actinomycetales; Corynebacteriaceae; Mycobacterium.
OC NCBI_TaxID=1773;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
  Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
  Badcock K., Bauman D., Brown D., Chillingworth T., Connor R.,
  Davies K., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
  Hornsby T., Jagels K., Krogh J., McLean J., Moule S., Murphy L.,
  Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
  Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
  Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
  complete genome sequence."
RL Nature 393:537-544(1998).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Ohkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
  Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
  Kolonay J.F., Nelson W.C., Umayan L.A., Ermolenko M.D., Salzberg S.L.,
  Bishai W.;
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
  Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
  laboratory strains."
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLPTHCEROL
  SYNTHESIS.
CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETINES (BY
  similarity).
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
```

```

CC -----
DR EMBL: Z74697; CAA38988.1; -.
DR EMBL: AE007122; AAK47328.1; -.
DR TIGR: MT3000; RV2931; -.
DR TubercuList; RV2931; -.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Pantane attach.
DR Pfam; PF00109; ketoacyl-synt 1.
DR Pfam; PF00550; pp-binding 2.
DR Pfam; PF00698; Acyl_transf; 1.
DR Pfam; PF02801; ketoacyl-synt C; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
DR PROSITE; PS00606; B KETOACYL SYNTHASE; 1.
DR PROSITE; PS50075; ACP DOMAIN; 2.
KM Multifunctional enzyme, Oxidoreductase; Transferase; NADP; Repeat;
  Phosphopantetheine; Complete proteome.
FT DOMAIN 7 80 ACYL CARRIER (ACP) 1.
FT NP_BIND 1764 1833 ACYL CARRIER (ACP) 2.
FT BINDING 62 62 NADP (POTENTIAL).
FT ACT_SITE 273 273 PHOSPHOPANTHETINE (POTENTIAL).
FT BINDING 1796 1796 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT CONFLICT 624 624 D -> E (IN REF. 2).
FT CONFLICT 877 877 R -> H (IN REF. 2).
FT CONFLICT 1323 1323 G -> S (IN REF. 2).
SQ SEQUENCE 1876 AA; 198834 MW; D9783DBD8792110 CRC64;

Query Match      22.8%; Score 217; DB 1; Length 1876;
Best Local Similarity 32.5%; Pred. No. 4.5e-11;
Matches 62; Conservative 25; Mismatches 86; Indels 18; Gaps 5;

QY 4 LINGTGALATHLTHLTHOPTHLLTSRTGHTPHR-----HAGHLTQO-----L 49
DB 1494 LITGGLGALGLMADWL-ADRGARLVLTRT-PLPPRDWQLDLDTELRRRIDAIRAL 1551
QY 50 QKGIHLITTCDSNPNQLOQLN-----NTTPQHPLTVVHTAGIILDDATLNTLPTQLN 107
DB 1552 EMGQVVEVAADVCGREVQALARDGDGAPRIGIHAAGITDQVLTSTGDAVRQ 1611
QY 108 VLRKASAHLLHQLTQHTPLTAVLYSSAATFGAPCOANYAANAVALAHHRHTH 167
DB 1612 VMPKXIGSGVLDAPPGPSVDFVLTASAAGIFGIPGQSYAANSYDALARARQOG 1671
QY 168 LPAISLAWGTW 178
DB 1672 CHTMSLDWVAM 1682

RESULT 8
PKSK_BACSU STANDARD; PRT; 4447 AA.
AC P40803;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DN Putative polyketide synthase pksk (PKS).
GN PKSK.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxID=1423;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=168 / PBI424;
RX MEDLINE=95219083; PubMed=7704258;
RA Albertini A.M., Caramori T., Scoffone F., Scotti C., Galizzi A.;
RT "Sequence around the 159 degree region of the Bacillus subtilis
  genome: the pks locus spans 33.6 kb."
RL Microbiology 141:289-309(1995).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
```

RA MEDLINE=98044033; PubMed=9384377;
 RA Kunze F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriello R., Bousquet L., Brans A., Braun M., Brinell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano N.J., Carter N.M.,
 RA Choi S.K., Codani J.J., Comercon I.F., Cummings N.J., Daniel R.A.,
 RA Danilov K.D., Devine K.M., Duesterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Fritze C., Fujita M., Furlan J., Fabbret C., Ferrari E., Foulger D.,
 RA Fritz S.Y., Gasser P., Goffeau A., Golligly E.J., Grandi G.,
 RA Gilchrist G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Gilbert H., Holmappell S., Hosono S., Hullo M.F., Itaya M., Jones U.,
 RA Jorle B., Karamela D., Kasehara Y., Kleier-Blanchard M., Klein C.,
 RA Kobayashi Y., Kottler P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine R.P., Liu H., Masuda S., Mauel N., Meligne C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesli D., Nakai S., Nodack M.,
 RA Noone D., O'Reilly M., Ogasawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Portolillo S., Prescott A.M.,
 RA Presencan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Solio B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 RA Tobar V., Uchiyama S., Vandenbol M., Vannier F., Vassart A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipst A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zanchin A.,
 RA "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RT Nature 390:249-256(1997).
 CC -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR
 CC THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN
 CC SECONDARY METABOLISM.
 CC -1- COPACTOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES
 CC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 5 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL, U11039, AAA85144.1; -;
 DR EMBL, Z99112, CAB13590.1; -;
 DR EMBL, Z99113, CAB13601.1; -;
 DR HSP, P14687, 1AMU.
 DR Subtilisin, BGI0930, pKex.
 DR Interpro: IPR000873; AMP-bind.
 DR Interpro: IPR001242; Condensacn.
 DR Interpro: IPR000794; Ketocacyl-synt.
 DR Interpro: IPR003880; Ppanent attach.
 DR Pfam: PF00109; Ketocacyl-synt; 3.
 DR Pfam: PF00501; AMP-binding; 1.
 DR Pfam: PF00550; PP-binding; 5.
 DR Pfam: PF00668; Condensation; 1.
 DR Pfam: PF02801; Ketocacyl-synt; 3.
 DR PRINTS, PR00154; AMPBINDING.
 DR PROSITE, PS00012; PHOSPHOPANTETHEINE; 3.
 DR PROSITE, PS00455; AMP-BINDING; 1.
 DR PROSITE, PS00606; B-KETOACYL SYNTHASE; 2.
 DR PROSITE, PS50075; ACP DOMAIN; 5.
 KM Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;
 KM Phosphopantetheine; Multifunctional enzyme; Repeat; Ligase;
 KM Complete proteome.
 FT DOMAIN 1 68 ACYL CARRIER (ACP) 1.
 FT DOMAIN 1063 1130 ACYL CARRIER (ACP) 2.

FT DOMAIN 2516 2589 ACYL CARRIER (ACP) 3.
 FT DOMAIN 2618 2687 ACYL CARRIER (ACP) 4.
 FT DOMAIN 2618 3937 ACYL CARRIER (ACP) 5.
 FT BINDING 31 31 PHOSPHOPANTETHEINE (POTENTIAL).
 FT BINDING 1093 1093 PHOSPHOPANTETHEINE (POTENTIAL).
 FT BINDING 2552 2552 PHOSPHOPANTETHEINE (POTENTIAL).
 FT BINDING 2650 2650 PHOSPHOPANTETHEINE (POTENTIAL).
 FT ACT SITE 2915 2915 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT BINDING 3900 3900 PHOSPHOPANTETHEINE (POTENTIAL).
 FT ACT SITE 4147 4147 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 SO SEQUENCE 4447 AA; 496058 MM; 96CDBA46C0A5C057 CXC64;
 Query Match 21.3%; Score 203; DB 1; Length 4447;
 Best Local Similarity 30.7%; Pred. No. 2,2e-09;
 Matches 61; Conservative 28; Mismatches 80; Indels 30; Gaps 7;
 QY 4 LITGCGALATHTLTHLTHPTQPHLLTSR--TGHTTPHAQLTT-----QL 49
 DB 3579 LITGCTRGIGGLCARFACGYVKLVLTGREQLPREBWARPKTNTSLAEXIOAVREL 3638
 QY 50 QOKGHLITTTCDTSNPDLQOLNTIIPQH-----PTTVHTAGILDDATLTN--T 101
 DB 3639 EAKGVQVEMSLTSLSDAOVEQTL-----OHKRTGPIGVCHGCLTMDTLAFIRKT 3693
 QY 102 PTOLNNVLRKAKASHALHLQLTQHTPLTAFLVYSSAAATFG--APCOANYAANAVALDAL 159
 DB 3694 SDDIQGVLEPKVSGTLTVRHVNEBLOFVLFSSVAILIPLSAQADVAMANSYMDYF 3753
 QY 160 AHHRHTHLPLPNTSIANGTW 178
 DB 3754 A-EAHOXKHAPIISVQPMNW 3771
 RESULT 9
 PAS_RAT STANDARD; PRT; 2505 AA.
 ID PAS_RAT 009187; 009190;
 AC P12785; Q64717; Q09187; Q09190;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [Includes: EC 2.3.1.38; EC 2.3.1.39;
 DE EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89240686; PubMed=2717611;
 RA Amy C.M., Witkowski A., Naggert J., Williams B., Randhawa Z.,
 RA Smith S.;
 RT "Molecular cloning and sequencing of cDNA encoding the entire rat
 RT fatty acid synthase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3114-3118(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=93075999; PubMed=1339331;
 RA Beck K.F., Schlegelmann R., Stachopoulos I., Klein H., Hoch J.,
 RA Schwaiger M.;
 RT "The fatty acid synthase (FAS) gene and its promoter in Rattus
 RT norvegicus.";
 RL DNA Seq. 21:359-386(1992).
 RN [3]
 RP SEQUENCE OF 75-2505 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Mammary gland;
 RX MEDLINE=89128431; PubMed=2915923;
 RA Schweitzer M., Takebayashi K., Beck K.F., Schlegelmann R.;
 RT "rat mammary gland fatty acid synthase: localization of the
 RT constituent domains and two functional polyadenylation/termination
 RT signals in the cDNA.";
 RL Nucleic Acids Res. 17:567-586(1989).

[4] SEQUENCE OF 2085-2505 FROM N.A.
RP TISSUE=Mammary gland;
RC MEDLINE=88087240; PubMed=2891707;
RA Naggett J., Witkowski A., Mikkelsen J., Smith S.;
RT "Molecular cloning and sequencing of a cDNA encoding the thioesterase
RT domain of the rat fatty acid synthetase.";
RL J. Biol. Chem. 263:1146-1150(1988).
RN [5]
RP SEQUENCE OF 1921-2324 FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=87246646; PubMed=3109907;
RA Witkowski A., Naggett J., Mikkelsen J., Smith S.;
RT "Molecular cloning and sequencing of a cDNA encoding the acyl carrier
RT protein and its flanking domains in the mammalian fatty acid
RT synthetase.";
RL Eur. J. Biochem. 165:601-606(1987).
RN
CC -1 FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
CC ACYL CARRIER PROTEIN.
CC -1 CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
CC long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).
CC -1 CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
CC acetyl-[acyl-carrier protein].
CC -1 CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
CC malonyl-[acyl-carrier protein].
CC -1 CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl]-
CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
CC [acyl-carrier protein].
CC -1 CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC -1 CATALYTIC ACTIVITY: (3R)-3-hydroxyheptadecanoyl-[acyl-carrier protein]
CC = 2-hexadecanoyl-[acyl-carrier protein] + H(2)O.
CC -1 CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
CC -1 CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl]-
CC carrier protein] + oleate.
CC -1 SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb.ch/announce/isb-slb.ch>).
CC or send an email to license@isb-slb.ch.

DR EMBL; M67672; AAA57219.1; -
DR EMBL; X62888; CAA44679.1; -
DR EMBL; X62889; CAA44680.1; -
DR EMBL; X13415; CAA31780.1; -
DR EMBL; X13527; CAA31882.1; -
DR EMBL; J03514; AAA41144.1; -
DR PIR; A30313; XIRTEA.
DR InterPro; IPR001272; AC_transferase.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Pantine_attach.
DR InterPro; IPR000051; SAM_bind.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00107; adh_zinc.1.
DR Pfam; PF00109; ketoacyl-synt.1.
DR Pfam; PF00550; pp-binding.1.
DR Pfam; PF00698; Acyl_transf.1.
DR Pfam; PF00975; Thioesterase.1.
DR Pfam; PF02801; ketoacyl-synt_C.1.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR PROSITE; PS00075; ACP_DOMAIN; 1.
KM Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
KM Hydroxylase; Oxidoreductase; Transferase; Lyase; NADP;
KM Pyridoxal phosphate.

FT DOMAIN 1 413
 FT DOMAIN 429 817
 FT DOMAIN 1629 1857
 FT DOMAIN 1858 2113
 FT DOMAIN 2118 2174
 FT DOMAIN 2202 2505
 FT ACT_SITE 161 161
 FT ACT_SITE 581 581
 FT ACT_SITE 1665 1682
 FT BINDING 1698 1698
 FT BINDING 1765 1780
 FT BINDING 2151 2151
 FT ACT_SITE 2302 2302
 FT ACT_SITE 2475 2475
 FT ACT_SITE 878 878
 FT CONFLICT 871 871
 FT CONFLICT 1967 1968
 FT CONFLICT 2085 2085
 FT CONFLICT 2106 2106
 FT CONFLICT 2296 2296
 SQ SEQUENCE 2505 AA; 272647 MW; 5810EC13037EF114 CRC64;
 Query March 21.1%; Score 201.5; DB 1; Length 2505;
 Best local Similarity 29.7%; Pred. No. 1.5e-09;
 Matches 52; Conservative 34; Mismatches 86; Indels 3; Gaps 2;
 QY 4 LITGTSALATLTHLTHLTHPOPTQHLHLSRTGPHPHQAHLTLQLOKGIHLTTTCDT 63
 Db 1882 IITGIGGFGFGLIARWLVL-RGAQRLVLTSSRGISRTGYQAKHVEWRRQGIHVLVSTSNV 1940
 QY 64 SNEPDLQQLANTIPRQHPPLTVVHTAGLIDATLTNTLTPOQLNNVLRKAKHSAILHLQLT 123
 Db 1941 SSLEGARALLAEATKGLPGVGVFNLAVLADMLENGTPELPFDVNRPKXNGTILMDRAT 2000
 QY 124 QHT--PLTATVLTSSAATFGAPQAVYAAANVYLDAHHRTHHHPATSIANG 176
 Db 2001 REACPFLDYVAFTSSVSCGRNMGQSYNGANSTMETICEQRHHDGLPLAVONG 2055
 RESULT 10
 PSB MYCTU STANDARD; PRT; 1538 AA.
 AC Q10978; O53234;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Polynucleotide synthetase polypeptide synthase PSB.
 GN PSB OR RV2932 OR MT3002 OR MTCY338.21 OR MTV011.01.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID:1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN:H37RV;
 RA MEDLINE=98295987; PubMed=9634230;
 RA Code S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gao S., Barry C.E. III, Tekala F.,
 RA Badcock K., Bauman D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagsle K., Krogh A., McLean J., Moutle S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultson J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland R., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gilm M.L., Hatt D., Hickey E.,

DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 5.
 DR PROSITE; PS00066; B-KETOACYL SYNTHASE; 1.
 DR PROSITE; PS50075; ACP DOMAIN; 5.
 KW Phosphopantetheine; Multifunctional enzyme; Repeat; Complete proteome.
 FT DOMAIN 211 280
 FT DOMAIN 382 759
 FT DOMAIN 937 1115
 FT DOMAIN 1409 1602
 FT DOMAIN 1687 1759
 FT DOMAIN 1876 2253
 FT DOMAIN 2491 2560
 FT DOMAIN 2632 2701
 FT DOMAIN 2823 3182
 FT DOMAIN 3575 3776
 FT DOMAIN 3854 3923
 FT BINDING 4019 4373
 FT BINDING 1723 1723
 FT BINDING 2523 2523
 FT BINDING 2664 2664
 FT BINDING 3886 3886
 SQ SEQUENCE 4427 AA; 493398 MW; 9612521E561AB9F2 CRC64;
 Query Match 20.5%; Score 195; DB 1; Length 4427;
 Best Local Similarity 28.6%; Pred. No. 1.1e-08;
 Matches 53; Conservative 29; Mismatches 95; Indels 8; Gaps 3;
 QY 1 GTTITGGTGAALHLLHLLTTHPTQHLITSTGTPPAQHLTQLOQKGIHLITTT 60
 DB 1410 GVVLTITGGAGGIGFIFATEILNQNDVAIVILGRS-PLDERKKKKLXALQDLGIAIYRQ 1468
 QY 61 CDTNPNPQLOOL-NTIPQPHLTVHTAGTLDATITNLTPQINNVLRKASAHLL 119
 DB 1469 ADLADKQTVADLKEIQNVGDDLGITISAGLIDKQNFIMKKKEEVQTVLAPKVAGLIHL 1528
 QY 120 HOLQHTPLTFAVLYSSAATFGAGQANVAANVYDALAHNR-----HTHLLPATSI 173
 DB 1529 DEARKDPIIDPFILFFSSGAGAVGAGQADYMANAFNAPFEYNQGAELHKRYGKTLSTV 1588
 QY 174 AMGTW 178
 DB 1589 CWPPLM 1593
 RESULT 12
 FAS-CHICK STANDARD; * PRT; 2511 AA.
 ID FAS-CHICK
 AC P12276;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [Inclides: EC 2.3.1.36; EC 2.3.1.39;
 EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN OR FAS.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-11.
 RC STRAIN=White leghorn; TISSUE=Liver;
 RX MEDLINE=95031085; PubMed=7944406;
 RA Huang W.-Y., Chirala S.S., Wakil S.J.;
 RT "Amino-terminal blocking group and sequence of the animal fatty acid
 synthase.";
 RL Arch. Biochem. Biophys. 314:45-49(1994).
 RN [2]
 RP SEQUENCE OF 75-1775 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89282777; PubMed=2734291;
 RA Holzer K.P., Liu W., Hammes G.G.;

RT "Molecular cloning and sequencing of chicken liver fatty acid
 RT synthase cDNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4387-4391(1989).
 RN [3]
 RP SEQUENCE OF 1568-2512 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=89139426; PubMed=2917973;
 RA Chirala S.S., Kasturi R., Pazirandeh M., Stolow D.T., Huang W.-Y.,
 RA Wakil S.J.;
 RT "A novel cDNA extension procedure. Isolation of chicken fatty acid
 RT synthase cDNA clones.";
 RL J. Biol. Chem. 264:3750-3757(1989).
 RN [4]
 RP SEQUENCE OF 1752-2512 FROM N.A.
 RX MEDLINE=88320436; PubMed=2842766;
 RA Yuan Z., Liu W., Hammes G.G.;
 RT "Molecular cloning and sequencing of DNA complementary to chicken
 RT liver fatty acid synthase mRNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:6328-6331(1988).
 RN [5]
 RP SEQUENCE OF 2202-2512 FROM N.A.
 RX MEDLINE=89088152; PubMed=3207710;
 RA Kasturi R., Chirala S.S., Pazirandeh M., Wakil S.J.;
 RT "Characterization of a genomic and cDNA clone coding for the
 RT thioesterase domain and 3' noncoding region of the chicken liver
 RT fatty acid synthase gene.";
 RL Biochemistry 27:7778-7785(1988).
 RN [6]
 RP SEQUENCE OF 2121-2209.
 RX MEDLINE=89192401; PubMed=2648999;
 RA Huang W.-Y., Stoops J.K., Wakil S.J.;
 RT "Complete amino acid sequence of chicken liver acyl carrier protein
 RT derived from the fatty acid synthase.";
 RL Arch. Biochem. Biophys. 270:92-98(1989).
 RN [7]
 RP SEQUENCE OF 2209-2508.
 RC STRAIN=White leghorn;
 RX MEDLINE=89088151; PubMed=3207709;
 RA Yang C.-Y., Huang W.-Y., Chirala S.S., Wakil S.J.;
 RT "Complete amino acid sequence of the thioesterase domain of chicken
 RT liver fatty acid synthase.";
 RL Biochemistry 27:7773-7777(1988).
 RN [8]
 RP SEQUENCE OF 667-674 AND 1698-1709.
 RX MEDLINE=89323081; PubMed=2751995;
 RA Chang S.I., Hammes G.G.;
 RT "Amino acid sequences of pyridoxal 5'-phosphate binding sites and
 RT fluorescence resonance energy transfer in chicken liver fatty acid
 RT synthase.";
 RL Biochemistry 28:3781-3788(1989).
 CC -1- FUNCTION. FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N-malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxypanthoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-
 CC carrier protein] + oleate.
 CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC CAUTION: REF.3 AND REF.5 SEQUENCES DIFFER FROM THAT SHOWN FROM
 CC POSITION 2351 ONWARDS DUE TO A FRAMESHIFT.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

CC -----
 CC EMBL: J04485; AAB46389.1; -
 CC EMBL: J03860; AAA48767.1; -
 CC EMBL: J02839; AAA82106.1; ALT_SEQ.
 CC PIR: A3318; XCHFA.
 CC PIR: A32015; A32015.
 CC InterPro: IPR001227; Ac transferase.
 CC InterPro: IPR02085; Adh zn family.
 CC InterPro: IPR000794; Ketocacyl-synt.
 CC InterPro: IPR003880; Prantne_attach.
 CC InterPro: IPR001031; Thioesterase.
 CC Pfam: PF00107; adh_zinc; 1.
 CC Pfam: PF00109; ketocacyl-synt; 1.
 CC Pfam: PF00550; pp-binding; 1.
 CC Pfam: PF00698; Acyl_transf; 1.
 CC Pfam: PF00975; Thioesterase; 1.
 CC Pfam: PF02801; ketocacyl-synt_C; 1.
 CC PROSITE: PS00012; PHOSPHOPANTHETHEINE; 1.
 CC PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.
 CC PROSITE: PS50075; ACP DOMAIN; 1.
 CC Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
 CC Transferrase; Hydrolyase; Oxidoreductase; Lyase; NADP; Acetylation;
 CC Alternative splicing; Pyridoxal phosphate.
 CC INIT MET 0
 CC DOMAIN 1 7411
 CC DOMAIN 427 815 BETA-KETOACYL SYNTHASE.
 CC DOMAIN 1638 1866 ACYL AND MALONYL TRANSFERASES.
 CC DOMAIN 1867 2119 ENOYL REDUCTASE.
 CC DOMAIN 2124 2180 BETA-KETOACYL REDUCTASE.
 CC DOMAIN 2209 2511 ACYL CARRIER (ACP).
 CC MOD RES 1 1 THIOESTERASE.
 CC ACT_SITE 160 160 ACETYLATION.
 CC ACT_SITE 579 579 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 CC ACT_SITE 579 579 BETA-KETOACYL TRANSFERASES (BY
 CC SIMILARITY).
 CC ACT_SITE 877 877 BETA-HYDROXYACYL DEHYDRATASE (BY
 CC SIMILARITY).
 CC NP BIND. 1674 1691 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC BINDING 1707 1707 NADP (R).
 CC NP BIND 1868 1903 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 CC BINDING 2157 2157 THIOESTERASE (BY SIMILARITY).
 CC ACT_SITE 2308 2308 THIOESTERASE (BY SIMILARITY).
 CC ACT_SITE 2481 2481 T -> TOCPSFLP (IN ISOFORM 1).
 CC VARSPPLIC 2348 2348 OL -> PV (IN REF. 2).
 CC CONFLICT 77 78 L -> A (IN REF. 2).
 CC CONFLICT 116 116 R -> S (IN REF. 2).
 CC CONFLICT 675 675 K -> N (IN REF. 2).
 CC CONFLICT 1169 1169 A -> T (IN REF. 2).
 CC CONFLICT 1178 1178 R -> H (IN REF. 2).
 CC CONFLICT 1191 1191 P -> L (IN REF. 2).
 CC CONFLICT 1198 1198 DN -> ND (IN REF. 2).
 CC CONFLICT 1286 1287 K -> E (IN REF. 2).
 CC CONFLICT 1372 1372 C -> Y (IN REF. 2).
 CC CONFLICT 1533 1533 W -> R (IN REF. 2).
 CC CONFLICT 1577 1577 OMAIALSMGC -> ASSHCHRLHGLA
 CC CONFLICT 1685 1696 (IN REF. 2; AAA48767).
 CC CONFLICT 1732 1732 O -> E (IN REF. 2).
 CC CONFLICT 1745 1745 S -> N (IN REF. 2).
 CC SEQUENCE 2511 AA; 274648 MW; 622039DACS15D3F CRC64;

Query Match 20.1%; Score 191.5; DB 1; Length 2511;
 Best Local Similarity 29.7%; Pred. No. 1,1e-08;
 Matches 52; Conservative 31; Mismatches 89; Indels 3; Gaps 3;
 4 LITGSGALNTLTHLTTQPTQHLLLSRTGPHTHAHLTTLOQKGINLTTTCDT 63

Db 1890 IITGGGCGFGLAQL-IERGAQKLVLSRSRGIRGYOAKVREKALGIQVLVSTDV 1948
 Qy 64 SNPDLOQLNTIRPQHPLTTVHTAGLIDDAFLTLTPQLNNVIRAK-AHSANILLHQL 122
 Db 1949 GTLEGQLILEALKGPVGIGINLAVLKDANIENQBELPVEVKNPKISGLHLDWT 2008
 Qy 123 TQHTP-LTAFVLVSSAAATFGAPQANYAANAVALDALAHRRTHHLPATSIAMG 176
 Db 2009 RKKCPDLDFVVSSTVSCGNGNAGSYNGFANGAMERICGRHHDGLPGLAVQWG 2063

RESULT 13
 ID_FAS HUMAN STANDARD; PRT; 2504 AA.
 AC P49327;
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-2001 (Rel. 40, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [includes: EC 2.3.1.38; EC 2.3.1.39;
 DE EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN OR FAS.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96004605; PubMed=7567999;
 RA Jayakumar A., Tai M.-R., Huang W.-Y., Al-Feel W., Hsu M.,
 RA Abu-Elheiga L., Chirra S.S., Wakil S.J.;
 RA "Human fatty acid synthase: properties and molecular cloning.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:8695-8699 (1995).
 RN [2]
 RP SEQUENCE OF 753-758 AND 1285-1297.
 RX MEDLINE=94294365; PubMed=8022791;
 RA Kuhajda F.P., Jenner K., Wood F.D., Hennigar R.A., Jacobs L.B.,
 RA Dick J.D., Pasternack G.R.;
 RA "Fatty acid synthase: a potential selective target for
 RT antineoplastic therapy.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6379-6383 (1994).
 CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl]-
 CC carrier protein = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein]
 CC = 2-hexadecenyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl]-
 CC carrier protein + oleate.
 CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC -1- TISSUE SPECIFICITY: PROMINENT EXPRESSION IN BRAIN, LUNG, AND
 CC LIVER.
 CC -1- MISCELLANEOUS: THE RELATIVELY LOW BETA-KETOACYL SYNTHASE ACTIVITY
 CC MAY BE ATTRIBUTABLE TO THE LOW 4'-PHOSPHOPANTHETHEINE CONTENT OF
 CC THE PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U26644; AAC50259.1; -
DR Genew; HGNC:3594; FASN.
DR MIM; 600212; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Ppanine_attach.
DR InterPro; IPR000051; SAM bind.
DR InterPro; IPR01031; Thioesterase.
DR Pfam; PF00107; adh_zinc; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF00550; pp-binding; 1.
DR Pfam; PF00698; Acyl_transf; 1.
DR Pfam; PF00975; Thioesterase; 1.
DR Pfam; PF02801; ketoacyl-synt C; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR PROSITE; PS50075; ACP_DOMAIN; 1.
KM Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
KW Hydroxylase; Oxidoreductase; Transferase; Lyase; NADP;
KW Pyridoxal phosphate.
FT DOMAIN 1 413 BETA-KETOACYL SYNTHASE.
FT DOMAIN 428 815 ACYL AND MALONYL TRANSFERASES.
FT DOMAIN 1630 1857 ENOYL REDUCTASE.
FT DOMAIN 1858 2113 BETA-KETOACYL REDUCTASE.
FT DOMAIN 2118 2174 ACYL CARRIER (ACP).
FT DOMAIN 2202 2504 THIOESTERASE.
FT ACT_SITE 161 161 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT ACT_SITE 580 580 MALONYLTRANSFERASE (BY SIMILARITY).
FT ACT_SITE 876 876 BETA-HYDROXYACYL DEHYDRATASE (BY SIMILARITY).
FT NP_BIND 1666 1683 NADP (BR).
FT BINDING 1699 1699 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT NP_BIND 1880 1895 NADP (KR).
FT BINDING 2151 2151 PHOSPHOPANTHETHEINE (BY SIMILARITY).
FT ACT_SITE 2302 2302 THIOESTERASE (BY SIMILARITY).
FT ACT_SITE 2475 2475 THIOESTERASE (BY SIMILARITY).
SQ SEQUENCE 2504 AA; 273100 MW; 8AAFA9809B23BDFA CRC64;
Query Match 19.2%; Score 182.5; DB 1; Length 2504;
Best Local Similarity 27.4%; Pred. No. 6.7e-08;
Matches 48; Conservative 36; Mismatches 88; Indels 3; Gaps 2;
QY 4 LINGGALATHTLTHLTTTQPTQPHLLTSRTGPHTPHAQHLTTQLOQKGIHLTTTCDT 63
DB 1882 IIAAGLGAGFGLLELAQWL-IQRGVQKLVLTSSGIRTVQAKQVRMRROGLQGVSTSN 1940
QY 64 SNPDQLOQLNTIPRPHLTIVITAGLIDATLTNLPTQIANVLAARAKASHLHQLT 123
DB 1941 SILEGAGGLAEAAQGLGVGVAVLVRDGLLENQTFEPFQVCKPKXSGTINLRLVT 2000
QY 124 QHT--PTAFVLYGSAATFGAPQANVAAVIALAHRRHTHHIPATSIANG 176
DB 2001 REACPELDYFVFSVSCGRNAGOSNYGFANSMERICRRHGGPLGLAVQMG 2055
AC P19096; STANDARD; PRT; 838 AA.
RESULT 14
FAS_MOUSE
ID FAS_MOUSE STANDARD; PRT; 838 AA.
AC P19096;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE Fatty acid synthase (EC 2.3.1.85) [includes: EC 2.3.1.38; EC 2.3.1.39;
DE EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14]
DE (Fragment).
GN FASN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Liver;
RX MEDLINE=89149781; Pubmed=2920037;
RA Paulauskis J.D., Sul H.S.;
RT "Structure of mouse fatty acid synthase mRNA. Identification of the
RT two NADPH binding sites".
RL Biochem. Biophys. Res. Commun. 158:690-695 (1989).
CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
CC ACYL CARRIER PROTEIN.
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
CC -1- long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
CC acetyl-[acyl-carrier protein].
CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
CC malonyl-[acyl-carrier protein].
CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
CC [acyl-carrier protein].
CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxypalmitoyl-[acyl-carrier protein]
CC = 2-tetradecenoyl-[acyl-carrier protein] + H(2)O.
CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-
CC carrier protein] + oleate.
CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; X13135; CAA31525.1; -
DR PIR; A32262; A32262.
DR MGD; MGI:95485; Fasn.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR003880; Ppanine_attach.
DR Pfam; PF00107; adh_zinc; 1.
DR Pfam; PF00550; pp-binding; 1.
DR Pfam; PF00975; Thioesterase; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; PARTIAL.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
DR PROSITE; PS50075; ACP_DOMAIN; 1.
KM Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
KW Hydroxylase; Oxidoreductase; Transferase; Lyase; NADP.
FT NON_TER 1 1
FT DOMAIN 451 507 ACYL CARRIER (ACP).
FT DOMAIN 535 838 THIOESTERASE.
FT BINDING 484 484 PHOSPHOPANTHETHEINE (BY SIMILARITY).
FT ACT_SITE 635 635 THIOESTERASE (BY SIMILARITY).
FT ACT_SITE 808 808 THIOESTERASE (BY SIMILARITY).
SQ SEQUENCE 838 AA; 91212 MW; 32C40E5CA5A8641 CRC64;
Query Match 18.8%; Score 179.5; DB 1; Length 838;
Best Local Similarity 29.5%; Pred. No. 3.2e-08;
Matches 52; Conservative 34; Mismatches 85; Indels 5; Gaps 4;
QY 4 LINGGALATHTLTHLTTTQPTQPHLLTSRTGPHTPHAQHLTTQLOQKGIHLTTTCDT 62
DB 215 IITGGAGFGLLELAQWL-RGQRLVLTSSGIRTVQAKH-REMRROGLQGVSTSN 272
QY 63 TSNPDQLOQLNTIPRPHLTIVITAGLIDATLTNLPTQIANVLAARAKASHLHQLT 122


```

DB 273 VSSLEGARALIAETKLPVGVFNLCGHRDMLNQLTPELPODVKPKYCNLNLMDRA 332
QY 123 TOHT--PLTAFVLYSSAAATFGAPGQANYAANAYLDALAHHRHTHLPTSTAMG 176
DB 333 TREACEBELDYFVAPSSVSCGRGNAGQNYCPANSTMERICEORRHVDVLPGLAOWMG 388

RESULT 15
PKSM_BACSU STANDARD; PRT; 4273 AA.
AC P40872; 031781;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative polyketide synthase pksM.
GN PKSM OR PksY.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1423;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; Pubmed=9384377;
RA Kunz F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Besiieres P., Bolotin A., Borchert S.,
RA Borriss R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruesch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.D., Daniel R.A.,
RA Denicot F., Devyne K.M., Dusterhoft A., Ertllich S.D., Emerson P.T.,
RA Ertlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grand G.,
RA Giuseppe G., Guy B.J., Hage K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Huilo M.F., Itaya M., Jones L.,
RA Jorle B., Karamata D., Kasehara Y., Kleier-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazerevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moserl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,
RA Paro V., Pohl T.M., Portelle D., Portwille S., Prescott A.M.,
RA Plescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serron S.J., Serron P., Shin B.S., Solio B.,
RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Takeo V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viart A., Wambutt R., Wedler E., Wedler H., Weitznegger T.,
RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RN Mature 390:249-256(1997).
[2]
RP SEQUENCE OF 1-1763 FROM N.A.
RC STRAIN=168 / PBL424;
RA Tognoni A., Grandi G.;
RA Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR
CC THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN
CC SECONDARY METABOLISM.
CC -!- COPACTOR: CONTAINS 4 COVALENTLY BOUND PHOSPHOPANTETHEINES
CC (POTENTIAL).
CC -!- SIMILARITY: CONTAINS 4 ACYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announce/)

```

```

CC or send an email to license@isb.ch.
CC -----
DR EMBL: Z99113; CAB3603.1; -
DR EMBL: Z35133; CA84505.1; -
DR Subtilin, BG10931; pksM.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR000794; ketoacyl-synt.
DR InterPro: IPR001601; Methyltransf.
DR InterPro: IPR003880; Ppantne attach.
DR InterPro: IPR000051; SAM_bind.
DR Pfam: PF00106; adh_short; 1.
DR Pfam: PF00109; ketoacyl-synt; 3.
DR Pfam: PF00550; pp-binding; 4.
DR Pfam: PF02801; ketoacyl-synt C; 3.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 2.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE: PS00075; ACP_DOMAIN; 4.
DR Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;
KW Phosphopantetheine; Multifunctional enzyme; Repeat; Complete proteome.
FT DOMAIN 295 364
FT DOMAIN 396 834
FT DOMAIN 1190 2258
FT DOMAIN 2322 2737
FT DOMAIN 3532 3947
FT DOMAIN 3410 3483
FT DOMAIN 4140 4209
FT BINDING 327 327
FT BINDING 2222 2222
FT ACT SITE 2476 2476
FT BINDING 3446 3446
FT ACT SITE 3690 3690
FT BINDING 4172 4172
FT CONFLICT 103 103
FT CONFLICT 276 276
FT CONFLICT 289 289
SQ SEQUENCE 4273 AA; 477459 MW; 3BBFCF1A250AEB5A CRC64;

Query Match 15.3%; Score 146; DB 1; Length 4273;
Best Local Similarity 26.7%; Pred. No. 0.0002;
Matches 48; Conservative 27; Mismatches 101; Indels 4; Gaps 4;

QY 1 GTTLTGTGATLATHLTHLTHPTQPHLLTSRTGPRPHAOHLTTLOQKGINHTTT 60
DB 3145 GTVLTIGVGGLGGLYLPKALKANY-ANLILTERS-PPNDEKQKQKELKLDGSEMYAE 3202
QY 61 CPTSNPDQLOQLNTTPQH-PLTYIHTAGILDDATLTNLTQUNVLRARASHAHL 119
DB 3203 ADVSDPIAMGDCVKGKGRFGAINGVHAAGIESDAIPDKIESFORIIEPKINGTIAL 3262
QY 120 HOLTOHTPLTAFVLYSSAAATFGAPGQANYAANAYLDALAHHRHTHLPT-STAMGTW 176
DB 3263 DEWLKNEDDLDFMVCYSSSAVALDFFCCDYAIGNFQMAIYAQYRNELHNGKTVINMPW 3322

```

Search completed: June 17, 2003, 13:02:35
 Job time : 2.94115 secs

THIS PAGE BLANK (USP10)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:55:32 ; Search time 3.97473 Seconds
(without alignments)
4305.183 Million cell updates/sec

Title: US-09-914-286-3_COPY_1715_1892

Perfect score: 953
Sequence: 1 GTTLITGTCGALATHLTHL.....LAHRRHTHLPATSIAMGTW 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	452	47.4	6260	2 T30228	polyketide synthase
2	449	47.1	8563	2 T30226	polyketide synthase
3	440	46.2	10223	2 T30225	polyketide synthase
4	439	46.1	4735	2 T17463	rifamycin polyketide
5	431	45.2	4151	2 G70944	polyketide synthase
6	429.5	45.1	6420	2 T30283	polyketide synthase
7	428.5	45.0	4613	2 T17409	polyketide synthase
8	421.5	44.2	2100	2 T03223	probable polyketide
9	419.5	44.0	7576	2 T17428	FK506 polyketide synthase
10	419	44.0	2124	2 T28658	polyketide synthase
11	415.5	43.6	2723	2 T03221	probable polyketide
12	411	43.1	3739	2 T17410	polyketide synthase
13	409.5	43.0	502	2 A70985	probable polyketide
14	400.5	42.0	1937	2 T03224	probable polyketide
15	392.5	41.2	1562	2 T17411	polyketide synthase
16	392.5	41.0	2162	2 T03222	probable polyketide
17	391	41.0	2126	2 H70621	probable polyketide
18	389.5	40.9	3491	2 T43331	probable 6-deoxyer
19	372.5	39.1	2103	2 G86925	probable polyketide
20	371.5	39.0	1616	2 G70668	polyketide synthase
21	367.5	38.6	3573	2 S23070	erythronolide synth
22	364.5	38.2	2518	2 A12140	polyketide synthase
23	351	36.8	3413	2 T17467	rifamycin polyketide
24	349.5	36.7	2478	2 AH2140	polyketide synthase
25	343	36.0	5069	2 T17464	rifamycin polyketide
26	334.5	35.1	1728	2 T17465	rifamycin polyketide
27	326	34.2	3172	2 S22012	erythronolide synth
28	312	32.7	1718	2 S13395	6-deoxyerythronolide
29	304.5	32.0	1763	2 T17465	rifamycin polyketide

30	303	31.8	1570	2 AC2012	hypothetical prote
31	300.5	31.5	1582	2 E70876	probable polyketide
32	295	31.0	3519	2 S43048	polyketide synthase
33	294.5	30.9	2126	2 E70522	probable polyketide
34	286	30.0	1198	2 T28678	polyketide synthase
35	282.5	29.6	2108	2 H70819	probable polyketide
36	278.5	29.2	2111	2 A70668	mycrocetate synth
37	271.5	28.5	2110	2 B44110	mycrocetate synth
38	270.5	28.4	2116	2 C86926	probable mycrocetate
39	247.5	26.0	2118	2 S72705	mycrocetate synth
40	233	24.4	958	2 S73012	polyketide synthase
41	231	24.4	1871	2 A87204	polyketide synthase
42	231.5	24.3	2188	2 A70984	probable polyketide
43	223	23.4	4930	2 E63679	polyketide synthet
44	220.5	23.1	1795	2 D97312	hypothetical prote
45	218.5	22.9	3161	2 T30342	protein HmP1 - Ye

ALIGNMENTS

RESULT 1

T30228
polyketide synthase - Streptomyces hygroscopicus
C/Species: Streptomyces hygroscopicus
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #ext_change 17-Nov-2000
C/Accession: T30228
R/Paratid, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Strau
Gene 169, 9-16, 1996
A/Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hy-
A/Reference number: Z20782; MUID:96186896; PMID:8635756
A/Accession: T30228
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-6260 <APA>
A/Cross-references: EMBL:X86780; NID:9987088; PID:9987102; PIDN:CAA60462.1
C/Genetics:
A/Genes: xpc
C/Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protei
C/Keywords: carrier protein
F/55-451/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F/1536-1607/Domain: acyl carrier protein homology <ACP1>
F/1651-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F/3123-3194/Domain: acyl carrier protein homology <ACP2>
F/3238-3632/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F/3725-3989/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F/5114-5185/Domain: acyl carrier protein homology <ACP3>
F/5242-5638/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F/5759-6044/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F/6135-6206/Domain: acyl carrier protein homology <ACP4>

Query Match 47.4%; Score 452; DB 2; Length 6260;

Best Local Similarity 52.8%; Pred. No. 6, 2e-31; Matches 94; Conservative 19; Mismatches 61; Indels 4; Gaps 1;

QY 1 GTTLITGTCGALATHLTHLTHQPTLHLSRTPRHNAOHLTQLOKSHLTIT 60
DB 4840 GTTLITGTCGALATHLTHLTHQPTLHLSRTPRHNAOHLTQLOKSHLTIT 4895
QY 61 CTSNPDQLOQLMTIPQHPRTTVITHTAGIIDDATLTMTPTQLNNVLRKASHSLH 120
DB 4896 CQSDRAGLAVLAVAGSPHPLTAIVHTAGALDDGVESLTQQLDTVLRPKADGAMH 4955
QY 121 QLTQHTPLTAFLVLYSSAAATFGAPGANYAANAYDALAHRRHTHLPATSIAMGTW 178
DB 4956 ELTQNTDLAFLVYSSAAATFGAPGANYAANAYDALAHRRHTHLPATSIAMGTW 5013

RESULT 2

T30226
polyketide synthase - Streptomyces hygroscopicus
C/Species: Streptomyces hygroscopicus
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #ext_change 01-Dec-2000

C:Accession: T30226
R:Rapriccio, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Staun
Gene 169, 9-16, 1996
A>Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hyg
A:Reference number: Z20782; MUID:96186896; PMID:8635756
A:Accession: T30226
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-8563 <APa>
A:Cross-references: EMBL:X86780; NID:9387088; PID:9387100; PIDN:CAA60460.1
C:Genetics:
A:Gene: rapA
C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:154-503/Domain: acetate-CoA ligase homology <AC1>
F:1189-1124/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F:1317-2091/Domain: 3-oxoacyl-[acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F:3139-3270/Domain: acyl carrier protein homology <ACP1>
F:3314-3706/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F:4787-4858/Domain: acyl carrier protein homology <ACP2>
F:4902-5293/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F:5386-5659/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F:6760-6831/Domain: acyl carrier protein homology <ACP3>
F:6875-7269/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F:7362-7636/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
F:8412-8483/Domain: acyl carrier protein homology <ACP4>
F:18822/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 47.1%; Score 449; DB 2; Length 8563;
Best Local Similarity 52.2%; Pred. No. 1.7e-30;
Matches 93; Conservative 21; Mismatches 60; Gaps 1;

```
QY 1 GTTTLTGSGALATLNTLNTLTHOFTONLLTSRGRNTPRANOHLTTOLOOKGNLTTT 60
Db 2931 GTTTLTGSGAGLIGLIRHLLAENHGRHLLLSRTTPD----QALIKELAEIGALVDAT 2986
QY 61 CDTSNPDQLQOLNMTIPRONPLTTHYHAGLIDDAATLNTLPTOYANPLRAKASHNLTH 120
Db 2987 CDVSRAGIARLAVLAVSPENPLTVAIHTHAGALDDCVSVSLTTOQDLYLRPADCAWHLH 3046
QY 121 QLTQHTPLTAFVLSSAATFCAFCQANVAAANALDLDLNNHRTNHLPATSIAMGTW 178
Db 3047 ELTQNTDILAAFMVSSAAGVLGSAQCGNYAANAANAVDLALAEORRGGELPALAVAAIGTW 3104
```

RESULT 3
T30225
polyketide synthase - *Streptomyces hygroscopicus*

C:\Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000

R. Aparicio, J. F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S. F.; Khaw, L. E.; Staunton, Gene 169, 9-16, 1996

A;Reference number: Z20782; MUID:96186896; PMID:8635756

A/Status: preliminary; translated from GB/EMBL/DBJ

A;Residues: 1-10223 <APA>

C;Genetics:

C/Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protein

F;54-449/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
E:1533-1504/Domain: acyl carrier protein homology <ACP1>

F;16447-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OA
F;2135-2408/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AM

E;3172-3243/Domains: acyl carrier protein homology <ACP2>
E;3287-3681/Domains: 3-oxoacyl-[acyl]-carrier-protein synthase I homology <OAS>

F:5160-5231/Domain: acyl carrier protein homology <ACP3>
F:3//8-4052/Domain: [acyl-carrier-protein] s-malonyltransferase homology <AM

F52/5-586//Domadl: 3-oxoacyl-lacyl-carrier-protein] вычислает 1 homology <0.454>

F_6776-6847//Domain: acyl carrier protein homology <ACP4>
F_6891-7285//Domain: 3-oxoacyl-(acyl-carrier-protein) synthase I homology <OAS5>
F_8411-8482//Domain: acyl carrier protein homology <ACS>
F_8556-8921//Domain: 3-oxoacyl-(acyl-carrier-protein) synthase I homology <OAS5>
F_9012-9265//Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
F_10069-10140//Domain: acyl carrier protein homology <ACP>

Query Match	46.2%	Score 440;	DB 2;	Length 10223;
Best Local Similarity	50.6%	Pred. No. 1,3e-29;		
Matches 90;	Conservative 23;	Mismatches 61;	Indels 4;	Gaps 1;

[illegible]

RESULT 4
T17463
rifamycin polyketide synthase modules 1-3 - Amycolatopsis mediterranei

C:\Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
C:\Accession: T17463
Rischupp, T.

A:Reference number: Z18802
A:Accession: T17463
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4735 <SCH>
A:Cross-references: EMBL:AJ223012; NID:e1227119; PID:e1227120; PIDD:CAA11035.1
A:Experimental source: strain LB5 AJ136
A:Superfamily: acyl carrier protein homology; acetate-CoA ligase homology
C:Keywords: carrier protein
E:53-500/Domain: acetate-CoA ligase homology <AC1>
E:543-610/Domain: acyl carrier protein homology <ACP1>

F:4578-4649/Domain: acyl carrier protein homology <ACPA>
Query Match 46.1%; Score 439; DB 2; Length 4735;

Matches 92; Conservative 18; Mismatches 64; Indels 4; Gaps 1;

QY 1 GTTLLTGTCGATATLTHLLTTHQPTOHLHLLTSRTGHTPNAQHLTLOLQXKIHLLITTT 60

[illegible]

Year	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

THE UNIVERSITY OF CHICAGO

[illegible]

RESULT 5
C70944

probable polyketide synthase RV2048c - *Mycobacterium tuberculosis* (strain H37RV)

C;Accession: G370944
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

R; Cole, S.T.; Davies, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, L.; Connor, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

kajandaream, m.a.; rogers, o.; kurtel, s.; seeger, n.; svelton, s.; squares, s.

Nature 393, 537-544, 1998
A/Authors: Squire, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Bartell, B.G.
A/Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A/Reference number: A70500; PMID:9825987; PMID:964230
A/Accession: G70944
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-4151 <COL>
A/Cross-references: GB:AL021899; GB:AL123456; NID:93424282; PIDN:CAA17262.1; PID:g289678
A/Experimental source: strain H37RV
C/Genetics:
A/Genes: fli2
C/Superfamily: *Mycobacterium tuberculosis* probable polyketide synthase Rv2048c; 3-oxoacyl
homology; [acyl-carrier-protein] S-malonyltransferase homology
C/Keywords: carrier protein
P/559-841/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
P/560-841/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
P/560-1861/Domain: short-chain alcohol dehydrogenase homology <SAD1>
P/563-2038/Domain: acyl carrier protein homology <ACP1>
P/2078-2478/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
P/2582-2860/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
P/3710-3891/Domain: short-chain alcohol dehydrogenase homology <SAD2>
P/3995-4066/Domain: acyl carrier protein homology <ACP2>

Query Match 45.2%; Score 431; DB 2; Length 4151;
Best Local Similarity 46.6%; Pred. No. 2,7e-29;
Matches 83; Conservative 32; Mismatches 63; Indels 0; Gaps 0;

QY 1 GTTLITGCGALATLTHLTHLTHQPTQHLTLRTGHTPHAOHLTLQLOKGINLTTT 60
DB 1680 GTVLTIGCGTGMAGSAARHVAHVGKRLVLSRRGDAGAAELVAELAAQAQVVA 1739
QY 61 CDTSPDQLOQLINTIPPOHPLTVITHTAGILDDATLTNLTPTQNNVLRKASHLH 120
DB 1740 CDAAADAAALAKVADIPVQHPILSGVHTACALDDAVVMSLTBRDVVLSKDAAMHLH 1799
QY 121 QLTQHTPLPAFLVYSSAAATFGAPGQANVAANAYLDALAHNRHTNHLPTSIANGTW 178
DB 1800 ELTRDLIDPAFLVYSSAAATFGAPGQANVAANAYLDALAHNRHTNHLPTSIANGTW 1857

RESULT 6
T30283
polyketide synthase - Streptomyces sp. (strain MA6548)
C/Species: Streptomyces sp.
A/Variety: strain MA6548
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
A/Accession: T30283
R/Notamed: H.; Cai, S.J.; Shafliee, S.J.; Elliston, K.O.
A/Title: Structural organization of a multifunctional polyketide synthase involved in th
A/Reference number: Z20806; PMID:97217427; PMID:9063448
A/Accession: T30283
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-6420 <MOT>
A/Cross-references: EMBL:Y10438; NID:e1014806; PID:e290681; PIDN:CAA71463.1
C/Genetics:
A/Note: fkbA
C/Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protei
C/Keywords: carrier protein
P/51-433/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
P/1930-2325/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
P/3391-3462/Domain: acyl carrier protein homology <ACP1>
P/3505-3900/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
P/3983-4254/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
P/5307-5378/Domain: acyl carrier protein homology <ACP2>
P/5431-5831/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
P/5932-6206/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
P/6293-6364/Domain: acyl carrier protein homology <ACP3>

Query Match 45.1%; Score 429.5; DB 2; Length 6420;
Best Local Similarity 54.3%; Pred. No. 6,2e-29;

Matches 95; Conservative 15; Mismatches 46; Indels 19; Gaps 4;

QY 4 LITGCGALATLTHLTHLTHQPTQHLTLRTGHTPHAOHLTLQLOKGINLTTT 63
DB 3140 LITGSGGTLAGIARL--NHRHTYLSSTRPPPTP-----GTH---PCDL 3182
QY 64 SMDQLOQLINTIPPOHPLTVITHTAGILDDATLTNLTPTQNNVLRKASHLH 123
DB 3183 TDPQTQALVTHIP--QPLTGFHTPAALTDATLTNLTPTQNNVLRKASHLH 3240
QY 124 QHTPLPAFLVYSSAAATFGAPGQANVAANAYLDALAHNRHTNHLPTSIANGTW 178
DB 3241 QNQPHTPLPAFLVYSSAAATFGAPGQANVAANAYLDALAHNRHTNHLPTSIANGTW 3295

RESULT 7
T17409
polyketide synthase type I - Streptomyces venezuelae
C/Species: Streptomyces venezuelae
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
A/Accession: T17409
R/Xue, Y.; Zhao, L.; Liu, H.W.; Sherman, D.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 12111-12116, 1998
A/Title: A gene cluster for macrolide antibiotic biosynthesis in streptomyces venezuela
A/Reference number: Z18773; PMID:98445333; PMID:9770448
A/Accession: T17409
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-4613 <XUE>
A/Cross-references: EMBL:AF079139; NID:93808326; PID:9380834; PIDN:AA69329.1
C/Genetics:
A/Genes: pikA
C/Superfamily: acyl carrier protein homology
C/Keywords: antibiotic biosynthesis; carrier protein
P/1010-1081/Domain: acyl carrier protein homology <ACP1>
P/2495-2566/Domain: acyl carrier protein homology <ACP2>
P/4407-4478/Domain: acyl carrier protein homology <ACP3>

Query Match 45.0%; Score 428.5; DB 2; Length 4613;
Best Local Similarity 48.9%; Pred. No. 5,1e-29;
Matches 87; Conservative 22; Mismatches 68; Indels 1; Gaps 1;

QY 1 GTTLITGCGALATLTHLTHLTHQPTQHLTLRTGHTPHAOHLTLQLOKGINLTTT 60
DB 2208 GTVLTIGCGALGSHAAARM--AHNGEHLIVSRSEQAPGATQLAELTASARVTIA 2266
QY 61 CDTSPDQLOQLINTIPPOHPLTVITHTAGILDDATLTNLTPTQNNVLRKASHLH 120
DB 2267 CDVADPHAKRTLLDAPETPLPAVVTHTAGALDDGIVTLTAEQVRAHRAKAVGASVLD 2326
QY 121 QLTQHTPLPAFLVYSSAAATFGAPGQANVAANAYLDALAHNRHTNHLPTSIANGTW 178
DB 2327 ELTRDLIDPAFLVYSSAAATFGAPGQANVAANAYLDALAHNRHTNHLPTSIANGTW 2384

RESULT 8
T03223
probable polyketide synthase module 3 - Streptomyces hygroscopicus
C/Species: Streptomyces hygroscopicus
C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 26-May-2000
A/Accession: T03223
R/Ruan, X.; Stassl, D.; Lax, S.; Katz, L.
Gene 203, 1-9, 1997
A/Title: A second type-I PKS gene cluster isolated from Streptomyces hygroscopicus ATCC
A/Reference number: Z14848; PMID:98085969; PMID:9426000
A/Accession: T03223
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2100 <RUAN>
A/Cross-references: EMBL:AF007101; NID:92624946; PIDN:AA638063.1; PID:g2624950
A/Experimental source: ATCC 29253
C/Superfamily: mycotoxin acid synthase; 3-oxoacyl-[acyl-carrier-protein] synthase I ho
nase homology; [acyl-carrier-protein] S-malonyltransferase homology

A;Reference number: Z18773; MUID:9845333; PMID:9770448
A;Accession: T17411
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1562 <XUE>
A;Cross-references: EMBL:AF079138; NID:G3808326; PID:G3800836; PIDN:AAC69331.1
C;Genetics:
A;Gene: plkAIII
C;Superfamily: Streptomyces hygrosopicus probable polyketide synthase module 4; 3-oxoad
homology; [acyl-carrier-protein] S-malonyltransferase homology
C;Keywords: antibiotic biosynthesis; carrier protein
F;1403-1474/Domain: acyl carrier protein homology <ACP>

Query Match 41.2%; Score 392.5; DB 2; Length 1562;
Best Local Similarity 46.6%; Pred. No. 2e-26;
Matches 83; Conservative 21; Mismatches 73; Indels 1; Gaps 1;

QY 1 GTTITGAGTALATLTHLTTTQPTQHLITSRTPHQAHLTTQLOQKGIHLTTT 60
||| ||| ||| ||| : ||| ||| ||| : ||| ||| ||| : |||
Db 1116 GTVLTITGAGTALGSHAARWM-AHHGAHLLVSRSGEQAPGATQLTAEITFASGARVTIAA 1174
||| ||| ||| ||| : ||| ||| ||| : ||| ||| ||| : |||

QY 61 CPTSNPDQLOQLNTIPQHELTIVITAGILDDATLTNLTPTQANVLPKASHAHLH 120
||| ||| ||| ||| : ||| ||| ||| : ||| ||| ||| : |||
Db 1175 CDVADPHAMRTLLDAIPETPLTAIVHTAGAGGDPDVTGPEDIARILGKTSGAELVD 1234
||| ||| ||| ||| : ||| ||| ||| : ||| ||| ||| : |||

QY 121 QLTQHTPLTAFLVLYSSAAATFGAPQANVAAANAYLDALAHRRTHHLPATSIAMGTW 178
||| ||| ||| ||| : ||| ||| ||| : ||| ||| ||| : |||
Db 1235 DLRGTFPLDAFVLVYSSNAGVWGSGSGVYAAANAHLDALARRRARGETATSVAMGLW 1292
||| ||| ||| ||| : ||| ||| ||| : ||| ||| ||| : |||

Search completed: June 17, 2003, 13:12:21
Job time : 3.97473 secs.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 13:08:02 / Search time 5.54613 Seconds
(without alignments)
3431.399 Million cell updates/sec

Title: US-09-914-286-3_COPY_1715_1892

Perfect score: 953
Sequence: 1 GTTLITGCGALATHTLTHL.....LAHRRHTHLPATSIAMGTW 178

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications_AA:
1: /cgn2_6/ptodaca/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodaca/1/pubpaa/PCR_NEW_PUB pep.*
3: /cgn2_6/ptodaca/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodaca/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodaca/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodaca/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodaca/1/pubpaa/PCRUS_PUBCOMB pep.*
8: /cgn2_6/ptodaca/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodaca/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodaca/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodaca/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodaca/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodaca/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodaca/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	428.5	45.0	4551	US-09-793-708-1	Sequence 1, Appl1
2	428.5	45.0	4613	US-09-860-846-31	Sequence 31, Appl1
3	428.5	45.0	4613	US-09-988-3848-31	Sequence 31, Appl1
4	428.5	45.0	4613	US-09-836-821-31	Sequence 31, Appl1
5	428.5	45.0	4613	US-09-861-289-31	Sequence 31, Appl1
6	428.5	45.0	11877	US-09-860-846-6	Sequence 6, Appl1
7	428.5	45.0	11877	US-09-836-821-6	Sequence 6, Appl1
8	428.5	45.0	11877	US-09-861-289-6	Sequence 6, Appl1
9	428.5	45.0	12199	US-09-988-3848-6	Sequence 6, Appl1
10	412	43.2	5215	US-09-860-846-2	Sequence 2, Appl1
11	412	43.2	5215	US-09-988-3848-2	Sequence 2, Appl1
12	412	43.2	5215	US-09-836-821-2	Sequence 2, Appl1
13	412	43.2	5215	US-09-861-289-2	Sequence 2, Appl1
14	411	43.1	3739	US-09-860-846-33	Sequence 33, Appl1
15	411	43.1	3739	US-09-988-3848-33	Sequence 33, Appl1
16	411	43.1	3739	US-09-836-821-33	Sequence 33, Appl1
17	411	43.1	3739	US-09-793-708-2	Sequence 2, Appl1
18	411	43.1	3739	US-09-861-289-33	Sequence 33, Appl1
19	400	42.0	4150	US-09-808-880-2	Sequence 2, Appl1

20	392.5	41.2	1562	9	US-09-860-846-35	Sequence 35, Appl1
21	392.5	41.2	1562	9	US-09-988-3848-35	Sequence 35, Appl1
22	392.5	41.2	1562	9	US-09-836-821-35	Sequence 35, Appl1
23	392.5	41.2	1562	9	US-09-793-708-33	Sequence 33, Appl1
24	392.5	41.2	1562	10	US-09-861-289-35	Sequence 35, Appl1
25	371.5	39.0	1616	9	US-09-712-363-262	Sequence 262, App
26	367.5	38.6	3816	9	US-09-808-880-3	Sequence 3, Appl1
27	362.5	38.0	7257	9	US-10-014-717-5	Sequence 5, Appl1
28	349	36.6	184	9	US-09-957-483-33	Sequence 33, Appl1
29	345	36.2	2439	9	US-10-014-717-7	Sequence 7, Appl1
30	337	35.2	1832	9	US-10-014-717-4	Sequence 4, Appl1
31	335	35.2	1834	9	US-09-957-483-35	Sequence 35, Appl1
32	328.5	34.5	3798	9	US-10-014-717-6	Sequence 6, Appl1
33	295	31.0	3519	9	US-09-808-880-4	Sequence 4, Appl1
34	263.5	27.6	2563	9	US-09-836-705-46	Sequence 46, Appl1
35	210	22.0	1837	9	US-09-712-363-261	Sequence 261, App
36	199	20.9	3032	9	US-09-836-705-44	Sequence 44, Appl1
37	181.5	19.0	2509	9	US-10-337-271-1	Sequence 1, Appl1
38	117	12.3	2756	9	US-10-331-061-7	Sequence 7, Appl1
39	100.5	10.5	711	9	US-10-124-800-18	Sequence 18, Appl1
40	100.5	10.5	1481	9	US-10-331-061-70	Sequence 70, Appl1
41	100.5	10.5	2910	9	US-10-124-800-2	Sequence 2, Appl1
42	98.5	10.3	244	10	US-09-815-242-10126	Sequence 10126, A
43	97.5	10.2	243	10	US-09-815-242-13360	Sequence 13360, A
44	97.5	10.2	243	10	US-09-815-242-13581	Sequence 13581, A
45	97.5	10.2	244	9	US-09-849-092-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-09-793-708-1
Sequence 1, Application US/09793708
Publication No. US200301045597A1
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002121
CURRENT APPLICATION NUMBER: US/09/793, 708
CURRENT FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: US 09/657,440
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: US 09/320, 878
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: US 09/141,908
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: US 09/073,538
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: US 08/846,247
PRIOR FILING DATE: 1997-04-30
PRIOR APPLICATION NUMBER: US 60/134,990
PRIOR FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 4551
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-793-708-1

Query Match 45.0%; Score 428.5; DB 9; Length 4551;

Best Local Similarity 48.9%; Pred. No. 1.1e-31;

Matches 87; Conservative 22; Mismatches 68; Indels 1; Gaps 1;

Ox: 1 GTTLITGCGALATHTLTHLTHPTOPQHLLTSTRTPTPHRPHQHTLTOLOKIKHILTTT 60
Db 2146 GTTLITGCGALGSHAAHWM-AHGAHLLVRSGEQAPGATVLTALTAAGARVTIAA 2204

[illegible]

```

RESULT 2
US-09-860-846-31
; Sequence 31, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 4613
; TYPE: prnt
; ORGANISM: Streptomyces venezuelae
; US-09-860-846-31

```

[illegible]

```

RESULT 3
US-09-988-384B-31
Sequence 31: Application US/093988384B
Publication No. US20030073824A1
GENERAL INFORMATION:
APPLICANT: Sherman, D. H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,536US1
CURRENT APPLICATION NUMBER: US/09/988,384B
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: PCT/US99/14398
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 53
SEQ ID NO 31
LENGTH: 4613
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-988-384B-31

```

Query Match	45.0%;	Score 428.5;	DB 9;	Length 4613;
Best Local Similarity	48.9%;	Pred. No. 1.1e-31;		
Matches	87;	Conservative	22;	Mismatches 68;
				Indels 1;
				Gaps 1;

Qy	Qy	Db
6	CDTENPOLCOOLNITRPGONHLETVNITAGLIDATLTNTPYNNVRYKASANDH	120
2267	CDVADPHAMETLIDATIPAEITPLTVAVTAGLIDGIVDTLTAEGVRAHRAKAGAVLD	23267
121	QIQTCHETLTVFYUSSAAATGEGASGNVYAAANVYDALNNHNTNNHLPATSIAMGTM	178
2327	BLTRDDLDLDAFVLFSSVSTLIGRGOSNVYARNHAYLDLAAKRAATGERSAVSVMGSM	2384

```

RESULT 4
US-09-836-821-31
Sequence 31, Application US/09836821
Publication NO. US20030087405A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and plitromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/836.821
CURRENT FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: 09/105,527
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31
LENGTH: 4613
TYPE: PR1
ORGANISM: Streptomyces venezuelae
US-09-836-821-31

```

[illegible]

```

RESULT 5 -31
US-09-861-289-31
; Sequence 31, Application US/09861289
; Patent No. US20020110897A1
GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438U51
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43

```

```
/ SOFTWARE: FaestSeq for Windows Version 3.0
/ SEQ ID NO 31
/ LENGTH: 4613
/ TYPE: PRF
/ ORGANISM: Streptomyces venezuelae
US-09-861-289-31

Query Match      45.0%; Score 428.5; DB 10; Length 4613;
Best Local Similarity 48.9%; Pred. No. 1.1e-31;
Matches 87; Conservative 22; Mismatches 68; Indels 1; Gaps 1;

QY 1 GTTLLTGCGALATLHNLHTTQPTQHLHLSRTGPTHTAHLTTQLOQKGIHLTTT 60
   |||||
DB 2208 GTVLTIGCGTALGSHARWM-AHHGAHLLVSRSGEAPGATQLTAEITASGARVTIAA 2266
   |||||

QY 61 CDTSNPDQLOQLNTIPRQHPLTTVTITAGILDDATLTNTLPTQLNNVLRKAKASHLH 120
   |||||
DB 2267 CDVADPHAKMTLLDAPAEPTPLTAHVHTAGALDDGIVDTLTAEQVRARAKAVGASVLD 2326
   |||||

QY 121 QLTQHTPLTAFLVYSSAAATFGAPGOANYAANAAYLDALAHHRHTHLPTATSIANGTW 178
   |||||
DB 2327 ELTRDLDLDAFVLFSSVSTLGIPOGQNVAPHNAAYLDALARRRATGRSAVSVAMGPW 2384
   |||||

RESULT 6
US-09-860-846-6
/ Sequence 6, Application US/09860846
/ Patent No. US20020164742A1
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, D.H.
/ APPLICANT: Liu, H.
/ APPLICANT: Xue, Y.
/ APPLICANT: Zhao, L.
/ TITLE OF INVENTION: DNA encoding methymycin and pikromycin
/ FILE REFERENCE: 600.438US1
/ CURRENT APPLICATION NUMBER: US/09/860.846
/ PRIOR FILING DATE: 2001-05-18
/ PRIOR APPLICATION NUMBER: 09/105.537
/ PRIOR FILING DATE: 1998-06-26
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FaestSeq for Windows Version 3.0
/ SEQ ID NO 6
/ LENGTH: 11877
/ TYPE: PRF
/ ORGANISM: Streptomyces venezuelae
US-09-860-846-6

Query Match      45.0%; Score 428.5; DB 9; Length 11877;
Best Local Similarity 48.9%; Pred. No. 3.9e-31;
Matches 87; Conservative 22; Mismatches 68; Indels 1; Gaps 1;

QY 1 GTTLLTGCGALATLHNLHTTQPTQHLHLSRTGPTHTAHLTTQLOQKGIHLTTT 60
   |||||
DB 2544 GTVLTIGCGTALGSHARWM-AHHGAHLLVSRSGEAPGATQLTAEITASGARVTIAA 2602
   |||||

QY 61 CDTSNPDQLOQLNTIPRQHPLTTVTITAGILDDATLTNTLPTQLNNVLRKAKASHLH 120
   |||||
DB 2603 CDVADPHAKMTLLDAPAEPTPLTAHVHTAGALDDGIVDTLTAEQVRARAKAVGASVLD 2662
   |||||

QY 121 QLTQHTPLTAFLVYSSAAATFGAPGOANYAANAAYLDALAHHRHTHLPTATSIANGTW 178
   |||||
DB 2663 ELTRDLDLDAFVLFSSVSTLGIPOGQNVAPHNAAYLDALARRRATGRSAVSVAMGPW 2720
   |||||

RESULT 7
US-09-836-821-6
/ Sequence 6, Application US/09836821
/ Publication No. US20030087405A1
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, D.H.
/ APPLICANT: Liu, H.
/ APPLICANT: Xue, Y.
/ APPLICANT: Zhao, L.
```

```
/ TITLE OF INVENTION: DNA encoding methymycin and pikromycin
/ FILE REFERENCE: 600.438US1
/ CURRENT APPLICATION NUMBER: US/09/836.821
/ CURRENT FILING DATE: 2001-04-17
/ PRIOR APPLICATION NUMBER: 09/105.537
/ PRIOR FILING DATE: 1998-06-26
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FaestSeq for Windows Version 3.0
/ SEQ ID NO 6
/ LENGTH: 11877
/ TYPE: PRF
/ ORGANISM: Streptomyces venezuelae
US-09-836-821-6

Query Match      45.0%; Score 428.5; DB 9; Length 11877;
Best Local Similarity 48.9%; Pred. No. 3.9e-31;
Matches 87; Conservative 22; Mismatches 68; Indels 1; Gaps 1;

QY 1 GTTLLTGCGALATLHNLHTTQPTQHLHLSRTGPTHTAHLTTQLOQKGIHLTTT 60
   |||||
DB 2544 GTVLTIGCGTALGSHARWM-AHHGAHLLVSRSGEAPGATQLTAEITASGARVTIAA 2602
   |||||

QY 61 CDTSNPDQLOQLNTIPRQHPLTTVTITAGILDDATLTNTLPTQLNNVLRKAKASHLH 120
   |||||
DB 2603 CDVADPHAKMTLLDAPAEPTPLTAHVHTAGALDDGIVDTLTAEQVRARAKAVGASVLD 2662
   |||||

QY 121 QLTQHTPLTAFLVYSSAAATFGAPGOANYAANAAYLDALAHHRHTHLPTATSIANGTW 178
   |||||
DB 2663 ELTRDLDLDAFVLFSSVSTLGIPOGQNVAPHNAAYLDALARRRATGRSAVSVAMGPW 2720
   |||||

RESULT 9
US-09-988-384B-6
/ Sequence 6, Application US/09861289
/ Patent No. US20020110897A1
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, D.H.
/ APPLICANT: Liu, H.
/ APPLICANT: Xue, Y.
/ APPLICANT: Zhao, L.
/ TITLE OF INVENTION: DNA encoding methymycin and pikromycin
/ FILE REFERENCE: 600.438US1
/ CURRENT APPLICATION NUMBER: US/09/861.289
/ CURRENT FILING DATE: 2001-05-18
/ PRIOR APPLICATION NUMBER: 09/105.537
/ PRIOR FILING DATE: 1998-06-26
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: FaestSeq for Windows Version 3.0
/ SEQ ID NO 6
/ LENGTH: 11877
/ TYPE: PRF
/ ORGANISM: Streptomyces venezuelae
US-09-861-289-6

Query Match      45.0%; Score 428.5; DB 10; Length 11877;
Best Local Similarity 48.9%; Pred. No. 3.9e-31;
Matches 87; Conservative 22; Mismatches 68; Indels 1; Gaps 1;

QY 1 GTTLLTGCGALATLHNLHTTQPTQHLHLSRTGPTHTAHLTTQLOQKGIHLTTT 60
   |||||
DB 2544 GTVLTIGCGTALGSHARWM-AHHGAHLLVSRSGEAPGATQLTAEITASGARVTIAA 2602
   |||||

QY 61 CDTSNPDQLOQLNTIPRQHPLTTVTITAGILDDATLTNTLPTQLNNVLRKAKASHLH 120
   |||||
DB 2603 CDVADPHAKMTLLDAPAEPTPLTAHVHTAGALDDGIVDTLTAEQVRARAKAVGASVLD 2662
   |||||

QY 121 QLTQHTPLTAFLVYSSAAATFGAPGOANYAANAAYLDALAHHRHTHLPTATSIANGTW 178
   |||||
DB 2663 ELTRDLDLDAFVLFSSVSTLGIPOGQNVAPHNAAYLDALARRRATGRSAVSVAMGPW 2720
   |||||
```

```

Sequence 6, Application US/09988384B
Publication No. US20030073824A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,536US1
CURRENT APPLICATION NUMBER: US/09/988,384B
PRIORITY FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: PCT/US99/14398
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 53
SEQ ID NO 6
LENGTH: 12199
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-988-384B-6

```

	Query Match	Best Local Similarity	Score	DB	Length
Matches	87	Conservative	22	Mismatches	66
				Indels	1
				Gaps	1
Qy	1	GTTLITGGTGAATLTHLHTTTHOPTHLLTSRTGPTPAOHLLTLOOKGIIHLITTT	60		
Db	2866	GTVLITGGTGAATGSHARWM-AHNGAEHLLVSRSGEQAFGATQLTAELTSGARVTIAA	2924		
Qy	61	CDTSNPDLQOLANTIPRPHLLTYITKAGILDDTLTNLPTQLNNVLAKASAILLH	120		
Db	2325	CDVADPHAMRLTLAIAPETPLTVVHTAGLIDGIVTLTAEQVRARAKAVASATLD	2984		
Qy	121	QLQPHPTLTAFLVSSAATFGAGQAAUAAANLYLDLAAHRRHNLPAISIMGTW	178		
Db	2385	ELTDDLDADAFLVSSVSTLTGTEQGSYAAPHNALDLAARRRATGSAVSVMGFW	3042		

```

RESULT 10
US-09-860-846-2
; Sequence 2, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438U1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-860-846-2

```

Query	Match	Similarity	43.2%	Score	412	DB	9	Length	5215
Beet	Local	Similarity	47.8%	Pred.	No.	5e-30			
Matches	85	Conservative	22	Mismatches	71	Indels	0	Gaps	0
Qy	1	GTTLITGGTALAHNLHTNLFTTHQPTOMLLTSRTPRTPHAQHLTTQLOQKGIHLTTT	60						
Db	2440	GTVLVTGGTGGALAHNLHVRHGRVRLLVSRGADAPAAADIGEDLAGIAGEVAFAA	24999						
Qy	61	CDTSNPDLQQLNTIPROPHLTVYIHAGILDDATLTNLTPTOLNNVLRKAKSAPHLH	120						
Db	2500	ADAAADRSILAPAITVPAHEPLTAVHTAGVADDAITVETALTPERIDAVLRKVDIANNH	25599						

QY 121 QLTHTPLTAVLYSSAATGARGOANYAAANAYLDALAHHRTHHLPTSIAWGTV 178
::: |||::: | ||||| ||||| |||: |||
Db 2560 ELTXDLRLDAFVLPFSVSIGIVTAGQANYAAANTGLDALLAHRAATGLAATSIAWGIWM 2617

```

RESULT 11
US-09-988-384B-2
; Sequence 2, Application US/09988384B
; Publication NO. US2003007382A1
GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-988-384B-2

```

Query Match	43.2%	Score 412; DB 9; Length 5215;
Best Local Similarity	47.8%	Pred. No. 5e-30;
Matches	85; Conservative 22; Mismatches 71; Indels 0; Gaps	
QY	1 GTTLLTGCGATGATLTHNLTHPHORPHMLLSRTPRHNAOHLTTOLOOKGHHLLTTT	608
DB	2440 GTVLVVGCGTGLGALLARHLVHNGVRLLVSRRAADRGAAIDGEDLAGAGAVAFAN	24
QY	61 CDTSNDPDLQOLLNTIPQPHPLTTVHTAGLDDATLNLPTTOLNNVLRAKASHHLH	12
DB	2500 ADAADRESLARAATATVPAAHEPLTAAVHTGAVDDATVEALTPERLDAVLPRKVDAAAMNH	25
QY	121 QLTQHPRLPRLFVLYSSAAATFGARPGQANNAANAYLDALNHRHTHNLPTATSIACWT	178
DB	2560 ELTKDRLDAFVLFSSVSGIVGAGQANNAANATGDAALAAHRAATGTAATSIANGLW	2617

```

RESULT 12
US-09-836-821-2
: Sequence 2, Application US/09836821
: Publication No. US20030087405A1
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600.438US1
: CURRENT APPLICATION NUMBER: US/09/836,821
: PRIORITY FILING DATE: 2001-04-17
: PRIORITY APPLICATION NUMBER: 09/105,537
: PRIORITY FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 5215
: TYPE: PRT
: ORGANISM: Streptomyces venezuelae
: US-09-836-821-2

```

Query Match 43.2%; Score 412; DB 9; Length 5215;
Best Local Similarity 47.8%; Pred. No. 5e-30;
Matches 85; Conservative 22; Mismatches 71; Indels 0; Gaps

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:54:32 ; Search time 3.23524 Seconds
(without alignments)
1618.821 Million cell updates/sec

Title: US-09-914-286-3_COPY_1715_1892

Perfect score: 953
Sequence: 1 GTTLITGTCGALATHTLTHL.....LAHRRHTHLPTSIAWGTV 178

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Issued_Patents_AA:*
1: /cgn2_6/ptodaca/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodaca/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodaca/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodaca/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodaca/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodaca/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	457	48.0	1864	2 US-08-804-227C-3	Sequence 3, Appl1
2	439	46.1	4545	2 US-08-804-227C-14	Sequence 14, Appl1
3	439	46.1	4550	2 US-08-804-227C-8	Sequence 8, Appl1
4	439	46.1	4550	2 US-08-804-198-2	Sequence 2, Appl1
5	436	45.8	5588	4 US-09-036-987A-6	Sequence 6, Appl1
6	436	45.8	5588	4 US-09-370-700-6	Sequence 6, Appl1
7	428.5	45.0	4551	3 US-09-320-878-1	Sequence 1, Appl1
8	428.5	45.0	4613	4 US-09-105-537-31	Sequence 31, Appl1
9	428.5	45.0	11877	4 US-09-105-537-6	Sequence 6, Appl1
10	423	44.4	1996	2 US-08-804-227C-9	Sequence 9, Appl1
11	423	44.4	1996	2 US-08-804-198-3	Sequence 3, Appl1
12	423	44.4	4472	2 US-08-804-227C-2	Sequence 2, Appl1
13	421	44.2	3724	2 US-08-804-227C-10	Sequence 10, Appl1
14	421	44.2	3724	2 US-08-804-198-4	Sequence 4, Appl1
15	413.5	43.4	6095	4 US-09-144-085-2	Sequence 2, Appl1
16	412	43.2	5215	4 US-09-105-537-2	Sequence 2, Appl1
17	411	43.1	3729	2 US-08-804-327C-4	Sequence 4, Appl1
18	411	43.1	3739	3 US-09-320-878-2	Sequence 2, Appl1
19	411	43.1	3739	3 US-09-105-537-33	Sequence 33, Appl1
20	400	42.0	4150	4 US-09-428-517-2	Sequence 2, Appl1
21	394.5	41.4	5087	4 US-09-144-085-1	Sequence 1, Appl1
22	392.5	41.2	1562	3 US-09-320-878-3	Sequence 3, Appl1
23	392.5	41.2	1562	4 US-09-105-537-35	Sequence 35, Appl1
24	391	41.0	3567	2 US-07-642-734C-4	Sequence 4, Appl1
25	391	41.0	3567	3 US-08-439-009A-4	Sequence 4, Appl1
26	389.5	40.9	3491	2 US-07-642-734C-2	Sequence 2, Appl1
27	389.5	40.9	3491	3 US-08-439-009A-2	Sequence 2, Appl1

28	389	40.8	2152	4 US-09-036-987A-3	Sequence 3, Appl1
29	389	40.8	2152	4 US-09-370-700-3	Sequence 3, Appl1
30	382.5	40.1	4928	4 US-09-036-987A-5	Sequence 5, Appl1
31	382.5	40.1	4928	4 US-09-370-700-5	Sequence 5, Appl1
32	373.5	39.2	1841	2 US-08-804-227C-6	Sequence 6, Appl1
33	367.5	38.6	3816	4 US-09-428-517-3	Sequence 3, Appl1
34	362.5	38.0	7257	3 US-09-335-409-5	Sequence 5, Appl1
35	362.5	38.0	7257	4 US-09-568-102-5	Sequence 5, Appl1
36	362.5	38.0	7257	4 US-09-567-969-5	Sequence 5, Appl1
37	362.5	38.0	7257	4 US-09-568-480-5	Sequence 5, Appl1
38	362.5	38.0	7257	4 US-09-568-486-5	Sequence 5, Appl1
39	362.5	38.0	7257	4 US-09-568-472-5	Sequence 5, Appl1
40	362.5	38.0	7257	4 US-09-567-899-5	Sequence 5, Appl1
41	355.5	37.3	2595	4 US-09-036-987A-2	Sequence 2, Appl1
42	355.5	37.3	2595	4 US-09-370-700-2	Sequence 2, Appl1
43	355	37.3	159	3 US-09-010-809-1	Sequence 1, Appl1
44	345	36.2	2259	4 US-09-413-814-70	Sequence 70, Appl1
45	345	36.2	2439	3 US-09-335-409-7	Sequence 7, Appl1

ALIGNMENTS

```
RESULT 1
US-08-804-227C-3
; Sequence 3, Application US/08804227C
; Patent No. 5876991
GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kunstos, Stuart A.
; APPLICANT: Roestek, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESS: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas G.
; REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1864 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-804-227C-3

Query Match 48.0%; Score 457; DB 2; Length 1864;
Best Local Similarity 50.6%; Pred. No. 3.3e-43;
Matches 90; Conservative 22; Mismatches 66; Indels 0; Gaps 0;

QY 1 GTTLITGTCGALATHTLTHLTHQPTQHLITRTGPHPHAGHLTTQLOKXIHLLITTT 60
DB 1414 GTTLITGTCGALATHTLTHLTHQPTQHLITRTGPHPHAGHLTTQLOKXIHLLITTT 1473
QY 61 CDTSNPDLOQLQINTIPQHPITTVITHTAGILDATLTNLTPTQLNNVLRKAKSHALLH 120
```

Db 1474 CDVADRALRALRLRVAGHPLTAVLHAGVLDGGLVLAQPPQGLAAVLARKADAVNHL 1533
QY 121 QLTQHTPLTAVLYSSAAATFGAGQANVYAAANVLDALHHRTTHLPATSIAMGTW 178
Db 1534 ELTQGHASRILYSSAGVLSAGQSGYAAANVLDSPAVWRSRGLPAVSLMGW 1591

RESULT 2

US-08-804-227C-14
Sequence 14, Application US/08804227C
Patent No. 5876991

GENERAL INFORMATION:
APPLICANT: DeHoff, Bradley S.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rosteck, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 4545 amino acids
TYPE: amino acid
TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-804-227C-14

Query Match 46.1%; Score 439; DB 2; Length 4545;

Best Local Similarity 49.4%; Pred. No. 1,4e-40;

Matches 88; Conservative 23; Mismatches 67; Indels 0; Gaps 0;

QY 1 GTTLITGTCGALATLHNLHTTHPTQHLITSTGTPHTPAOHLTTOLOOKGILHTTT 60
Db 4080 GTVLITGTCGALGRVASHLARRYGVRHMLVSRGPDAPBAGLERELAGLTATFLA 4139
QY 61 CDTSNPDQLOQLNTIPQHPHTTVIHTAGLIDDAITLNTLPQNLNNVLRKASHLH 120
Db 4140 CDLTIDIAVRKAAVAAPSDHPLTGVTHTAGVLDGALGTLRQRLDTVLPRKADAVNHL 4199
QY 121 QLTQHTPLTAVLYSSAAATFGAGQANVYAAANVLDALHHRTTHLPATSIAMGTW 178
Db 4200 EATLDRPLRAFLVLSAAAGLGRPGQASYAANAVLDALGARRAAGLPVAVSLMGIM 4257

RESULT 3

US-08-804-227C-8
Sequence 8, Application US/08804227C
Patent No. 5876991

GENERAL INFORMATION:
APPLICANT: DeHoff, Bradley S.
APPLICANT: Kuhstoss, Stuart A.

APPLICANT: Rosteck, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 4550 amino acids
TYPE: amino acid
TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-804-227C-8

Query Match 46.1%; Score 439; DB 2; Length 4550;

Best Local Similarity 49.4%; Pred. No. 1,4e-40;

Matches 88; Conservative 23; Mismatches 67; Indels 0; Gaps 0;

QY 1 GTTLITGTCGALATLHNLHTTHPTQHLITSTGTPHTPAOHLTTOLOOKGILHTTT 60
Db 4085 GTVLITGTCGALGRVASHLARRYGVRHMLVSRGPDAPBAGLERELAGLTATFLA 4144
QY 61 CDTSNPDQLOQLNTIPQHPHTTVIHTAGLIDDAITLNTLPQNLNNVLRKASHLH 120
Db 4145 CDLTIDIAVRKAAVAAPSDHPLTGVTHTAGVLDGALGTLRQRLDTVLPRKADAVNHL 4204
QY 121 QLTQHTPLTAVLYSSAAATFGAGQANVYAAANVLDALHHRTTHLPATSIAMGTW 178
Db 4205 EATLDRPLRAFLVLSAAAGLGRPGQASYAANAVLDALGARRAAGLPVAVSLMGIM 4262

US-08-804-198-2
Sequence 2, Application US/08804198
Patent No. 5945320

GENERAL INFORMATION:

APPLICANT: Bursgett, Stanley G.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rao, Nagaraja R.
APPLICANT: Richardson, Mark A.
APPLICANT: Rosteck, Paul R., Jr.
TITLE OF INVENTION: PLATENOLOIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh

OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4550 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-2

Query Match 46.1%; Score 439; DB 2; Length 4550;
Best Local Similarity 49.4%; Pred. No. 1,4e-40;
Matches 88; Conservative 23; Mismatches 67; Indels 0; Gaps 0;

QY 1 GTTLITGGTALATHLTHLTHQPTQHLLTSRTGPHPHAOHLTTQLOOKGHLTTT 60
DB 4085 GTVLITGGTALGRVSHLARRYGVHMLVSRGDAPEAGPLERLAGLVTATFLA 4144
QY 61 CDTSNPDLOQLNTIPPOHPLTVHTAGLDDATLTNLTPTQLNNVLRKASHAHLH 120
DB 4145 CDLTDEAVKAAVAASDHPHTGVTAGLDDGALTGLTROLDTVLPRKDAVNLH 4204
QY 121 QLTQHTPLTAFVLVSSAAATFGAPGOANYAANAAYDALAHHRHTHLPTSIAWGTW 178
DB 4205 EATLDRPLRAFVLPSAAGLGRGQASYAANAAYDALAGRAAGLPAVSLAMGLM 4262

RESULT 5

US-09-036-987A-6
Sequence 6, Application US/09036987A
Patent No. 6143526
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Meilio, Donald J.
APPLICANT: Treadway, Patci J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608

TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 5588 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-6

Query Match 45.8%; Score 436; DB 4; Length 5588;
Best Local Similarity 47.2%; Pred. No. 4.2e-40;
Matches 84; Conservative 31; Mismatches 63; Indels 0; Gaps 0;

QY 1 GTTLITGGTALATHLTHLTHQPTQHLLTSRTGPHPHAOHLTTQLOOKGHLTTT 60
DB 4933 GTVLITGGTGGALARRHLSAHGVHLLSRRLAAGADELVAELQAGADAVVA 4992
QY 61 CDTSNPDLOQLNTIPPOHPLTVHTAGLDDATLTNLTPTQLNNVLRKASHAHLH 120
DB 4993 CDSADDSLARLVAASVPAENPLRVVHAAGVLDGVLMSRPERLDAVLRPKVDAAWYLLH 5052
QY 121 QLTQHTPLTAFVLVSSAAATFGAPGOANYAANAAYDALAHHRHTHLPTSIAWGTW 178
DB 5053 ELTRELGSAFVLFSVAGLFGAGQSNYAAGNAFIDALAHCHQAOGLPALSLASGLM 5110

RESULT 6

US-09-370-700-6
Sequence 6, Application US/09370700
Patent No. 6274350
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patci J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 Div1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 5588
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-6

Query Match 45.8%; Score 436; DB 4; Length 5588;
Best Local Similarity 47.2%; Pred. No. 4.2e-40;
Matches 84; Conservative 31; Mismatches 63; Indels 0; Gaps 0;

QY 1 GTTLITGGTALATHLTHLTHQPTQHLLTSRTGPHPHAOHLTTQLOOKGHLTTT 60
DB 4933 GTVLITGGTGGALARRHLSAHGVHLLSRRLAAGADELVAELQAGADAVVA 4992
QY 61 CDTSNPDLOQLNTIPPOHPLTVHTAGLDDATLTNLTPTQLNNVLRKASHAHLH 120
DB 4993 CDSADDSLARLVAASVPAENPLRVVHAAGVLDGVLMSRPERLDAVLRPKVDAAWYLLH 5052
QY 121 QLTQHTPLTAFVLVSSAAATFGAPGOANYAANAAYDALAHHRHTHLPTSIAWGTW 178
DB 5053 ELTRELGSAFVLFSVAGLFGAGQSNYAAGNAFIDALAHCHQAOGLPALSLASGLM 5110

RESULT 7

US-09-320-878-1

```

Sequence 1 Application US/09320878A
Patent No. 6117659
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002120
CURRENT APPLICATION NUMBER: US/09/320,878A
CURRENT FILING DATE: 1999-05-27
EARLIER APPLICATION NUMBER: CIP OF 09/141,908
EARLIER FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: CIP OF 09/073,538
EARLIER FILING DATE: 1998-05-06
EARLIER APPLICATION NUMBER: CIP OF 08/846,247
EARLIER FILING DATE: 1997-04-30
EARLIER APPLICATION NUMBER: 60/119,139
EARLIER FILING DATE: 1999-02-08
EARLIER APPLICATION NUMBER: 60/100,880
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: 60/087,080
EARLIER FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 4551
TYPE: prt
ORGANISM: Streptomyces venezuelae
US-09-320-878-1

```

```

RESULT 8
US-09-105-537-31
: Sequence 31, Application US/09105537A
: Patent No. 6265202
: GENERAL INFORMATION:
: APPLICANT: Sherman, D. H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OR INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600.438US1
: CURRENT APPLICATION NUMBER: US/09/105.537A
: CURRENT FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 31
:
: LENGTH: 4613
:
: TYPE: PRT
:
: ORGANISM: Streptomyces venezuelae
US-09-105-537-31

45.0%; Score 428.5; DB 4; Length 4613;
Best Local Similarity 48.9%; Pred. No. 2,3e-39;
Matches 87; Conservative 22; Mismatches 66; Indels 1; Gaps 1;

```

```

QY      1 GTLLITGGTGAAMLNHLNLTTPQNLILNLTSGTHTNHAQNLPTLOOKGKHLITTT 60
Db      2208 GTVLITGGTGAAGSHAARMN-AHHGAHNLILNLTSSGQABGATQTLTALPSAGRAVTIAA 2266

QY      61 CDTNPPOLOOLNLTIPROHPLTVTVIRTAGLIDATLTNLTPTQNLNVNLRKAKASHILH 120
Db      2267 CDVADPRAMKRLTLDATPAETPLTTRVATTAAGLDDGIVDTLTAEQVRANRPAKAVGASVLD 2326

QY      121 QLTQHTLETLTFAVLSSAAATFGAPGQANVAANAAYLDALAHNRHTNHLRPTSIANGTW 178
Db      2327 ELTRDLDLDAFVLFSSTVSTLGIPIGQSNVAPNAUYLDALAAKRAATGRSAVSAMGFW 2384

```

```

RESULT 9
US-09-105-537-6
: Sequence 6, Application US/09105537A
: Patent No. 6265202
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600.438US1
: CURRENT APPLICATION NUMBER: US/09/105,537A
: CURRENT FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 11877
: TYPE: PRT
: ORGANISM: Streptomyces venezuelae
US-09-105-537-6

```

Query Match 45.0%; Score 428.5; DB 4; Length 11877;
 Best Local Similarity 48.9%; Pred. No. 8.9e-39;
 Matches 87; Conservative 22; Mismatches 68; Indels 1; Gaps 1;

```

QY      1 GTTLITGGTALATHLTHLHTTHOPTHLLLTSTGTGHTPHAOHLYTQLOOKGIHLITTT 60
Db      2544 GTVLITGGTATGSHARRWV-AHHGAEHLITVRSRSGQADGATQLTAEILTAASGARVTIAA 2602
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY      61 CDTISNPQOOLNTITPPQHPITVITAGLIDDATITNTITPQUNNVLAQAKASALILH 120
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      2603 CDVADPFAMKTLTDAPETAPPLVAVHTAAGLDGIVDTLTAAEVRRHRAKAAAGAVLD 2662
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY      121 QLTQHTPLTFLVLYSSAAATGAPGOANYAAUNYDALAHHRHTHLPATSIAMGFW 178
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      2663 ELTRDELIDLVAFSSVSTLGIPEGQGVNYPHNAYIDALAAARRATGRSAVSAMGFW 2720
        |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
  
```

RESULT 10
 US-08-804-227C-9
 Sequence 9, Application US/08804227C
 Patent No. 5876921
 GENERAL INFORMATION:
 APPLICANT: Dehoff, Bradley S.
 APPLICANT: Kuhstoss, Stuart A.
 APPLICANT: Rosteck, Paul R., Jr.
 APPLICANT: Sutton, Kimberly L.
 TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES.
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: THOMAS G. PLANT 1501
 STREET: LILLY CORPORATE CENTER
 CITY: INDIANAPOLIS
 STATE: IN
 COUNTRY: USA
 ZIP: 46285
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1996 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-227C-9

Query Match 44.4%; Score 423; DB 2; Length 1996;
Best Local Similarity 46.6%; Pred. No. 2.9e-39;
Matches 83; Conservative 26; Mismatches 69; Indels 0; Gaps 0;

QY 1 GTTLITGCGALATLTHLTHPTQPHLLTSRTGPHPHAOHLTTOLOOKGHLTTT 60
DB 1525 GTVLITGCGALAEFARHLVDRHKVRHLLVGRGPDAGVDRLVAELTESGAEVAVRA 1584
QY 61 CDTSNPDQLOQLNTIPPOHPLTVHTAGLIDATLTNLTPTQNNVLRKASHLHL 120
DB 1585 CDVTRDRLRLDLALDEHPLTCVHTAGLIDGVLSAQTAEIRIDVLRPKDAVAHLD 1644
QY 121 QLTQHTPLTFAVLSSAAATFGARQANVAAANAYLDALAHNRHTHLPATSIAMGTW 178
DB 1645 ELTREIGRVPLVLYSSVSATLSAGAGYAAANAFMDALARRCAAGHPALSLGGMW 1702

RESULT 11

US-08-804-198-3
Sequence 3, Application US/08804198
Patent No. 5945320
GENERAL INFORMATION:
APPLICANT: Burgett, Stanley G.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rao, Nagendra R.
APPLICANT: Richardson, Mark A.
APPLICANT: Kosteck, Paul R., Jr.
TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 1996 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-3

Query Match 44.4%; Score 423; DB 2; Length 1996;
Best Local Similarity 46.6%; Pred. No. 2.9e-39;
Matches 83; Conservative 26; Mismatches 69; Indels 0; Gaps 0;

QY 1 GTTLITGCGALATLTHLTHPTQPHLLTSRTGPHPHAOHLTTOLOOKGHLTTT 60
DB 1525 GTVLITGCGALAEFARHLVDRHKVRHLLVGRGPDAGVDRLVAELTESGAEVAVRA 1584
QY 61 CDTSNPDQLOQLNTIPPOHPLTVHTAGLIDATLTNLTPTQNNVLRKASHLHL 120
DB 1585 CDVTRDRLRLDLALDEHPLTCVHTAGLIDGVLSAQTAEIRIDVLRPKDAVAHLD 1644
QY 121 QLTQHTPLTFAVLSSAAATFGARQANVAAANAYLDALAHNRHTHLPATSIAMGTW 178
DB 1645 ELTREIGRVPLVLYSSVSATLSAGAGYAAANAFMDALARRCAAGHPALSLGGMW 1702

RESULT 12

US-08-804-227C-2
Sequence 2, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Suction, Kimberly L.
APPLICANT: Suction, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4472 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-804-227C-2

Query Match 44.4%; Score 423; DB 2; Length 4472;
Best Local Similarity 47.2%; Pred. No. 9.3e-39;
Matches 84; Conservative 22; Mismatches 72; Indels 0; Gaps 0;

QY 1 GTTLITGCGALATLTHLTHPTQPHLLTSRTGPHPHAOHLTTOLOOKGHLTTT 60
DB 4017 GTVLITGATGTGRLRLARHLAEHGVRLHLLSRGAAEGADELAELGAEBCFPA 4076
QY 61 CDTSNPDQLOQLNTIPPOHPLTVHTAGLIDATLTNLTPTQNNVLRKASHLHL 120

; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-144-085-2

Query Match 43.4%; Score 413.5; DB 4; Length 6095;
Best Local Similarity 48.0%; Pred. No. 1.8e-37;
Matches 86; Conservative 25; Mismatches 67; Indels 1; Gaps 1;
Qy 1 GTTLITGCGATLTHLTHHTTHQPTOHLTLTSRTGPHTPHACHLTTOLOOKGIH-LTIT 59
Db 1431 GTVLVTGCTGELGAIAMHLVRAGVRLVLTSTRGLRPGARELVOSLEELGAEVTVVA 1490
Qy 60 TCDTSNPDLQQLINTIPQHPITTVIHTAGILDATLTNLTPQLNNTLRKAKHSAHL 119
Db 1491 ACDSKREBVARVLGIDARPLSAVLHLGVLDDGVLTQTAERLSRVLA PKVDGALHL 1550
Qy 120 HQLQHTPLTAPVLYSSAATFGAPGOANTAAANAYLDALAHRRHTHLPTSTIANGTW 178
Db 1551 HELTRELIDLAFLVLFSSAAGTFGAAGOSNYAAANTPLDALAHRRGGGLAATSLANGFW 1609

Search completed: June 17, 2003, 13:09:27
Job time : 4.23524 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:50:57 / Search time 8.71975 Seconds
(without alignments)
2720.102 Million cell updates/sec

Title: US-09-914-286-3_COPY_1715_1892

Perfect score: 953

Sequence: 1 GTTLITGTCGATLATHLH.....LAHHRTHLPAISIAWGTW 178

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

A_Geneseq_101002:.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	953	100.0	3972	21	AA23749
2	953	100.0	3972	22	AA23749
3	953	100.0	3972	22	AA23749
4	948	99.5	6239	21	AA23750
5	948	99.5	6239	22	AA23750
6	500	52.5	5435	22	AA23750
7	483	50.7	11096	22	AA23750
8	475	49.8	2066	22	AA23750
9	461.5	48.4	9477	22	AA23750
10	457	48.0	1864	18	AA23750

11	456.5	47.9	6797	22	AA23758
12	452	47.4	4881	21	AA23751
13	452	47.4	4881	22	AA23751
14	441.5	46.3	7058	22	AA23751
15	439	46.1	4545	18	AA22611
16	439	46.1	4550	18	AA22611
17	439	46.1	4550	18	AA22611
18	439	46.1	4550	18	AA22611
19	436	45.8	5588	20	AA23930
20	436	45.8	5588	20	AA23930
21	429.5	45.1	1557	21	AA23930
22	429.5	45.1	1557	21	AA23930
23	429.5	45.1	1557	21	AA23930
24	429.5	45.1	1557	21	AA23930
25	429.5	45.1	1557	21	AA23930
26	428.5	45.0	4551	21	AA23930
27	428.5	45.0	4551	21	AA23930
28	428.5	45.0	4551	21	AA23930
29	428.5	45.0	4551	21	AA23930
30	428.5	45.0	4551	21	AA23930
31	425	44.6	3192	22	AA23930
32	423	44.4	1996	18	AA22607
33	423	44.4	1996	18	AA22607
34	423	44.4	1996	18	AA22607
35	421	44.2	3724	18	AA22607
36	421	44.2	3724	18	AA22607
37	415.5	43.6	3546	22	AA22607
38	415	43.5	4630	18	AA22607
39	413.5	43.4	6095	22	AA22607
40	412	43.2	4630	21	AA22607
41	411	43.1	3729	18	AA22603
42	411	43.1	3729	21	AA22603
43	411	43.1	3729	21	AA22603
44	411	43.1	3729	21	AA22603
45	411	43.1	3729	21	AA22603

ALIGNMENTS

RESULT 1
AA23749 standard; Protein: 3972 AA.
XX
AC AA23749;
XX
DT 10-JAN-2001 (first entry)
XX
DE S. avermectilis avermectin aglycon synthase protein SEQ ID NO:3.
XX
KW Streptomyces avermectilis; avermectin aglycon synthase; biosynthesis;
KW multifunctional enzyme; polyketide; avermectin; veterinary drug;
KW agrochemical.
XX
OS Streptomyces avermectilis.
XX
PN MO200050605-A1.
XX
PD 31-AUG-2000.
XX
PF 23-FEB-2000, 2000MO-JP1041.
XX
PR 24-FEB-1999, 99JP-0046961.
XX
(KITA) KITASATO INST.
XX
PI Omura S, Ikeda H;
XX
DR WPI, 2000-565458/52.
XX
DR N-PSDB; AAA2301.
XX
PT Avermectin aglycone synthase DNA and proteins encoded by all or part of it for the production of avermectin and its derivatives for drug and

PT agrochemical use -
XX
PS Claim 32; Page 203-223; 314pp; Japanese.
XX
CC The present sequence represents an avermectin aglycon synthase protein.
CC Also described are: (1) polypeptides encoded by all or part of the DNA;
CC (2) expression vectors containing the DNA; (3) host cells transformed by
CC the vectors; (4) preparation of the polypeptides by culture of the
CC transformants; (5) preparation of avermectin aglycon or its derivatives
CC by culture of transformed avermectin-producing microorganisms; and (6)
CC oligonucleotides of 5-60 bases in length containing sense or antisense
CC sequences from the avermectin aglycon synthase DNA. The enzymes are
CC useful for the production of modified forms of avermectin and of the
CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
CC and agrochemicals.
SQ Sequence 3972 AA;
Query Match 100.0%; Score 953; DB 21; Length 3972;
Best Local Similarity 100.0%; Pred. No. 1.7e-93;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTTLLTGATGALATHTLTHLTHPTOPHLLTSRTGPHRPAOHLTTLOOKGIHLTTT 60
DB 1715 GTTLLTGATGALATHTLTHLTHPTOPHLLTSRTGPHRPAOHLTTLOOKGIHLTTT 1774
QY 61 CDTSNPDQLOQLNTIPPOHPLTTVHTAGILDATLTNLTPTQLNNVLRKAKSAHLH 120
DB 1775 CDTSNPDQLOQLNTIPPOHPLTTVHTAGILDATLTNLTPTQLNNVLRKAKSAHLH 1834
QY 121 QLTQHTPLTAFLVYSSAAATFGAPGOANYAANAAYDALAHNRHTHLPTSIAWGTV 178
DB 1835 QLTQHTPLTAFLVYSSAAATFGAPGOANYAANAAYDALAHNRHTHLPTSIAWGTV 1892
RESULT 2
AAG65264
ID AAG65264 standard; Protein; 3972 AA.
XX
AC AAG65264;
XX
DT 04-DEC-2001 (first entry)
XX
DE Streptomycetes avermilitilis protein SEQ ID NO: 4.
XX
KM Avermectin aglycone synthase; AAS; avermectin derivative;
KM drug production; veterinary drug; pesticide.
XX
OS Streptomycetes avermilitilis.
XX
PN WO200162939-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-JP01381.
XX
PR 24-FEB-2000; 2000JP-0047405.
XX
PA (KYOWA) KYOWA HAKKO KOGYO KK.
PA (KITA) KITASATO INST.
XX
PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
XX
DR WPI; 2001-582053/65.
DR N-PSDB; AAH79277.
XX
PT New modified avermectin aglycone synthase derived from Streptomycetes
PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
PT drugs and pesticides -
XX
PS Claim 4; Page 167-180; 257pp; Japanese.
XX
CC The present invention relates to the production of modified derivatives

CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
CC avermilitilis. The activity of an acyl carrier protein (ACP),
CC beta-ketoadyl carrier protein synthase (KS), acyltransferase (AT),
CC beta-ketoadyl carrier protein reductase (KR), dehydratase (DH), enoyl
CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC suppressed. The process can be used in the production of drugs, veterinary
CC drugs and pesticides. The present sequence is an S. avermilitilis protein.
SQ Sequence 3972 AA;
Query Match 100.0%; Score 953; DB 22; Length 3972;
Best Local Similarity 100.0%; Pred. No. 1.7e-93;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTTLLTGATGALATHTLTHLTHPTOPHLLTSRTGPHRPAOHLTTLOOKGIHLTTT 60
DB 1715 GTTLLTGATGALATHTLTHLTHPTOPHLLTSRTGPHRPAOHLTTLOOKGIHLTTT 1774
QY 61 CDTSNPDQLOQLNTIPPOHPLTTVHTAGILDATLTNLTPTQLNNVLRKAKSAHLH 120
DB 1775 CDTSNPDQLOQLNTIPPOHPLTTVHTAGILDATLTNLTPTQLNNVLRKAKSAHLH 1834
QY 121 QLTQHTPLTAFLVYSSAAATFGAPGOANYAANAAYDALAHNRHTHLPTSIAWGTV 178
DB 1835 QLTQHTPLTAFLVYSSAAATFGAPGOANYAANAAYDALAHNRHTHLPTSIAWGTV 1892
RESULT 3
AAG65268
ID AAG65268 standard; Protein; 3972 AA.
XX
AC AAG65268;
XX
DT 04-DEC-2001 (first entry)
XX
DE Streptomycetes avermilitilis protein derivative SEQ ID NO: 8.
XX
KM Avermectin aglycone synthase; AAS; avermectin derivative;
KM drug production; veterinary drug; pesticide.
XX
OS Synthetic.
XX
PN WO200162939-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-JP01381.
XX
PR 24-FEB-2000; 2000JP-0047405.
XX
PA (KYOWA) KYOWA HAKKO KOGYO KK.
PA (KITA) KITASATO INST.
XX
PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
XX
DR WPI; 2001-582053/65.
DR N-PSDB; AAH79279.
XX
PT New modified avermectin aglycone synthase derived from Streptomycetes
PT avermectilis used in production of 22,23-dihydroavermectin B1a used in
PT drugs and pesticides -
XX
PS Claim 5; Page 235-248; 257pp; Japanese.
XX
CC The present invention relates to the production of modified derivatives
CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
CC avermilitilis. The activity of an acyl carrier protein (ACP),
CC beta-ketoadyl carrier protein synthase (KS), acyltransferase (AT),
CC beta-ketoadyl carrier protein reductase (KR), dehydratase (DH), enoyl
CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
CC suppressed. The process can be used in the production of drugs, veterinary
CC drugs and pesticides. The present sequence is derived from an S.
CC avermilitilis protein.

XX Sequence 3972 AA;
 Query Match 100.0%; Score 953; DB 22; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTLTGTCGATLTHLTHLTHPTQPHLLTSRTGPHPHAOHLTTLOOKGHLTTT 60
 DB 1715 GTTLTGTCGATLTHLTHLTHPTQPHLLTSRTGPHPHAOHLTTLOOKGHLTTT 1774

QY 61 CDTSNPDQLOQLNTIIPQHPPLTVITHTAGILDDATLTNLTPTQLNNVLRKAKHSAHLH 120
 DB 1775 CDTSNPDQLOQLNTIIPQHPPLTVITHTAGILDDATLTNLTPTQLNNVLRKAKHSAHLH 1834

QY 121 QLTQHPPLTAFVLYSSAAATFGAPGOANYAANAAYDALAHNRHTHLPTATSIANGTW 178
 DB 1835 QLTQHPPLTAFVLYSSAAATFGAPGOANYAANAAYDALAHNRHTHLPTATSIANGTW 1892

RESULT 4
 AAB23750 standard; Protein; 6239 AA.
 ID AAB23750;
 AC AAB23750;
 XX
 DT 10-JAN-2001 (first entry)
 XX
 DE 8. avermectin aglycon synthase protein SEQ ID NO:4.
 XX
 KM Streptomyces avermectilis; avermectin aglycon synthase; bioynthesis;
 KM multifunctional enzyme; polyketide; avermectin; veterinary drug;
 KM agrochemical.
 XX
 OS Streptomyces avermectilis.
 XX
 PN WC200050605-A1.
 XX
 PD 31-AUG-2000.
 XX
 PR 23-FEB-2000; 2000MO-JP01041.
 XX
 PR 24-FEB-1999; 99JP-0046961.
 XX
 PA (KITA) KITASATO INST.
 XX
 PI Omura S, Ikeda H;
 DR WPI; 2000-565458/52.
 DR N-PSDB; AAA92301.
 XX
 PT Avermectin aglycone synthase DNA and proteins encoded by all or part of
 PT it for the production of avermectin and its derivatives for drug and
 PT agrochemical use -
 XX
 PS Claim 32; Page 223-254; 314pp; Japanese.
 XX
 CC The present sequence represents an avermectin aglycon synthase protein.
 CC Also described are: (1) polypeptides encoded by all or part of the DNA;
 CC (2) expression vectors containing the DNA; (3) host cells transformed by
 CC the vectors; (4) preparation of the polypeptides by culture of the
 CC transformants; (5) preparation of avermectin aglycon or its derivatives
 CC by culture of transformed avermectin-producing microorganisms; and (6)
 CC oligonucleotides of 5-60 bases in length containing sense or antisense
 CC sequences from the avermectin aglycon synthase DNA. The enzymes are
 CC useful for the production of modified forms of avermectin and of the
 CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
 CC and agrochemicals.
 CC
 XX Sequence 6239 AA;
 Query Match 99.5%; Score 948; DB 21; Length 6239;
 Best Local Similarity 99.4%; Pred. No. 1.1e-92;

Matches 177; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTLTGTCGATLTHLTHLTHPTQPHLLTSRTGPHPHAOHLTTLOOKGHLTTT 60
 DB 2306 GTTLTGTCGATLTHLTHLTHPTQPHLLTSRTGPHPHAOHLTTLOOKGHLTTT 2365

QY 61 CDTSNPDQLOQLNTIIPQHPPLTVITHTAGILDDATLTNLTPTQLNNVLRKAKHSAHLH 120
 DB 2366 CDTSNPDQLOQLNTIIPQHPPLTVITHTAGILDDATLTNLTPTQLNNVLRKAKHSAHLH 2425

QY 121 QLTQHPPLTAFVLYSSAAATFGAPGOANYAANAAYDALAHNRHTHLPTATSIANGTW 178
 DB 2426 QLTQHPPLTAFVLYSSAAATFGAPGOANYAANAAYDALAHNRHTHLPTATSIANGTW 2483

RESULT 5
 AAG65265 standard; Protein; 6239 AA.
 ID AAG65265;
 AC AAG65265;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Streptomyces avermectilis protein SEQ ID NO: 5.
 XX
 KM Avermectin aglycone synthase; AAS; avermectin derivative;
 KM drug production; veterinary drug; pesticide.
 XX
 OS Streptomyces avermectilis.
 XX
 PN WC200162939-A1.
 XX
 PD 30-AUG-2001.
 XX
 PR 23-FEB-2001; 2001MO-JP01381.
 XX
 PR 24-FEB-2000; 2000JP-0047405.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 PA (KITA) KITASATO INST.
 XX
 PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
 DR WPI; 2001-582053/65.
 DR N-PSDB; AAH92277.
 XX
 PT New modified avermectin aglycone synthase derived from Streptomyces
 PT avermectilis used in production of 22,23-dihydrovermectin B1a used in
 PT drugs and pesticides -
 XX
 PS Claim 4; Page 180-201; 257pp; Japanese.
 XX
 CC The present invention relates to the production of modified derivatives
 CC of avermectin aglycone synthase (AAS) derived from Streptomyces
 CC avermectilis. The activity of an acyl carrier protein (ACP),
 CC beta-keetoacyl carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
 CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is an S. avermectilis protein.
 CC
 XX Sequence 6239 AA;
 Query Match 99.5%; Score 948; DB 22; Length 6239;
 Best Local Similarity 99.4%; Pred. No. 1.1e-92;
 Matches 177; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTTLTGTCGATLTHLTHLTHPTQPHLLTSRTGPHPHAOHLTTLOOKGHLTTT 60
 DB 2306 GTTLTGTCGATLTHLTHLTHPTQPHLLTSRTGPHPHAOHLTTLOOKGHLTTT 2365

QY 61 CDTSNPDQLOQLNTIIPQHPPLTVITHTAGILDDATLTNLTPTQLNNVLRKAKHSAHLH 120

FT	Domain	6166..6478	/label= AT8	/note= "Acyltransferase (AT) domain"
FT	Domain	6492..6704	/label= DH6_domain	/note= "Dehydratase (DH) domain"
FT	Domain	7038..7281	/label= KR6_domain	/note= "Ketoreductase (KR) domain"
FT	Domain	7315..7388	/label= ACP6_domain	/note= "Acyl carrier protein (ACP) domain"
FT	Domain	7408..7831	/label= KS7_domain	/note= "Ketosynthase (KS) domain"
FT	Domain	7939..8253	/label= AT7_domain	/note= "Acyltransferase (AT) domain"
FT	Domain	8267..8470	/label= DH7_domain	/note= "Dehydratase (DH) domain"
FT	Domain	8812..9086	/label= KR7_domain	/note= "Ketoreductase (KR) domain"
FT	Domain	9120..9193	/label= ACP7_domain	/note= "Acyl carrier protein (ACP) domain"
FT	Domain	9214..9637	/label= KS8_domain	/note= "Ketosynthase (KS) domain"
FT	Domain	9758..10072	/label= AT8_domain	/note= "Acyltransferase (AT) domain"
FT	Domain	10066..10289	/label= DH8_domain	/note= "Dehydratase (DH) domain"
FT	Domain	10657..10904	/label= KR8_domain	/note= "Ketoreductase (KR) domain"
FT	Domain	10939..11012	/label= ACP8_domain	/note= "Acyl carrier protein (ACP) domain"
FT	Domain	16-AUG-2001.		
XX	16-FEB-2001; 2001W0-GB00509.			
XX	08-FEB-2000; 2000GB-0002840.			
XX	10-APR-2000; 2000GB-0008786.			
XX	14-APR-2000; 2000GB-0009387.			
XX	(UYNO-) UNIV NORGE'S TEKNIKS NATUREVITENSKABELIGE.			
XX	(SNTF) SINTEF STIFTELSEN IND TEK FORSK.			
XX	(ALPH-) ALPHARMA AS.			
XX	(SINV-) SINVENT AS.			
XX	(DZIE-) DZIEGLEWSKA H.			
XX	(ZOTC-) ZOTCHEV S B.			
XX	(SEKU-) SEKUROVA O N.			
XX	(FJAE-) FJAEVYIK E.			
XX	(BRAU-) BRAUTASET T.			
XX	(STROM-) STROM A R.			
XX	Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR,			
XX	Valla S, Ellingsen TE, Sletta H, Gulliksen O,			
XX	WPI, 2001-557614/62.			
XX	N-PSDB, AAD17184.			
XX	New nystatin polyketide synthase polynucleotides and polypeptides			
XX	useful as antibiotic and antifungal -			

PS	Claim 15, Page 170-176; 266p; English.
XX	The present invention relates to the cloning and sequencing of the gene
CC	cluster encoding a modular type I polyketide synthase (PKS) enzyme
CC	involved in the biosynthesis of the macrocyclic antibiotic nystatin.
CC	The nystatin PKS is useful as antifungal antibiotics. The present
CC	sequence is a PKS type I encoding Streptomyces noursei nystatin
CC	gene, NysC.
XX	
XX	Sequence 11096 AA;
XX	
XX	Query Match 50.7%; Score 483; DB 22; Length 11096;
XX	Best Local Similarity 54.5%; Pred. NO. 9.5e-42; Indels 0; Gaps 0;
XX	Matches 97; Conservative 19; Mismatches 62;
Qy	1 GTTLLTGTGATLALNLTNHLTTTQPTQNLRLSRGTGPTNPAOHLTTLOOKGILNLTIT 60
Db	1381 GTVLTITGTGGLGAVLARIHLVATNGVYHLLARSRGPAADGADLTALTLGLATGATVIAA 1440
Qy	61 CDTSNPDOLQOLLNTTPRPHLTFTVITHTAGILDATLTNLTPQLNVLRAKASHALLH 120
Db	1441 CDVADPALADLTGLTVAPAGHPLTGVVHTAGVDDGVLSITLPORLDTVLPRKADAHLH 1500
Qy	121 QLTQHTLTLPVLYSSAAATFGARGQANVAAAYIDALAHNHTHTNHLPTSTANGTW 178
Db	1501 EATRHLDLDFVLFSSVAATLGSFGQANVAAAFIDALARRAATGLPATSTLAMPW 1558
RESULT 8	
AAE10146	
ID	AAE10146 standard; Protein; 2066 AA.
XX	
XX	AAE10146;
XX	
DT	29-NOV-2001 (first entry)
XX	
DE	Streptomyces noursei nystatin gene, NysC.
XX	
KW	Polyketide synthase; PKS; macrocyclic; nystatin; PKS gene cluster;
XX	antifungal; antibiotic; PKS type I.
OS	Streptomyces noursei.
XX	
XX	Key Location/Qualifiers
FT	34..457
FT	/label= KS18_domain
FT	/note= "ketosynthase (KS) domain"
FT	568..881
FT	/label= AT18_domain
FT	/note= "acyltransferase (AT) domain"
FT	898..1102
FT	/label= DH18_domain
FT	/note= "dehydratase (DH) domain"
FT	1416..1663
FT	/label= KR18_domain
FT	/note= "ketoreductase (KR) domain"
FT	1695..1769
FT	/label= ACP18_domain
FT	/note= "acyl carrier protein (ACP) domain"
FT	1849..2066
FT	/label= TE_domain
FT	/note= "thioesterase (TE) domain"
XX	
XX	WO200159126-A2.
XX	
XX	16-AUG-2001.
XX	
XX	08-FEB-2001; 2001WO-GB00509.
XX	
XX	08-FEB-2000; 2000GB-0002840.
XX	10-APR-2000; 2000GB-0008786.
XX	14-APR-2000; 2000GB-0009387.

```

PA (UNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
PA (SNTF) SINTER STIFTELSEN IND TEK FORSK.
PA (ALPH-) ALPHARMA AS.
PA (SINV-) SINVENT AS.
PA (DZIE/) DZIEGLEWSKA H.
PA (ZOTC/) ZOTCHEV S B.
PA (SEKU/) SEKUROVA O N.
PA (FJAE/) FJAEVIRK E.
PA (BRAU/) BRAUTASET T.
PA (STRO/) STROM A R.
XX
XX
XX Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;
PI Valia S, Ellingsen TE, Sletta H, Gulliksen O;
XX
XX WPI: 2001-557614/62.
XX N-PSDB; AADI7186.
DR
DR N-PSDB; AADI7186.
XX
XX
XX New nystatin polyketide synthase polynucleotides and polypeptides,
PT useful as antibiotics and antifungals -
XX
XX
XX Claim 15; Page 263-264; 266pp; English.
PS
XX
XX The present invention relates to the cloning and sequencing of the gene
CC cluster encoding a modular type I polyketide synthase (PKS) enzyme
CC involved in the biosynthesis of the macrocyclic antibiotic nystatin.
CC The nystatin PKS is useful as antifungal antibiotics. The present
CC sequence is a PKS type I encoding Streptomyces noursei nystatin gene,
CC NySk.
XX
XX
XX Sequence 2066 AA;
SQ
XX
XX
XX Query Match 49.8%; Score 475; DB 22; Length 2066;
XX Best Local Similarity 52.8%; Pred. No. 6,5e-42;
XX Matches 94; Conservative 17; Mismatches 67; Indels 0; Gaps 0;
QY
QY 1 GTTLITGAGALATHTLTHLTHQPTQHLTLTSRTGPHTPAHLTTLOQKGIHLTTT 60
DB 1416 GTVLITGAGTGLGAALRHILVTVGARRHLLAGRGEPAGELVAELTFOGADYVAA 1475
QY 61 CDTSNPQLOOLNTITPPQHLTVITAGLIDATITNTPTPOLANVLRKASAHLLH 120
DB 1476 CDVGRTRPALDALTVPAPHLTVAVHTAGLDLALGSLPPDALTATVLRKDAAMHLH 1535
QY 121 QLTQHTPLTAFVLSSAAATFGAPQANVAAANVYLDALHRRHTHLPTSTIMGTW 178
DB 1536 DATRGLDLAGFVLVSSVSGVLGSPQGNVAAANVYLDALHRRADQGLPLSLAMGPW 1593
XX
XX
XX RESULT 9
XX AAE10144
XX ID AAE10144 standard; Protein; 9477 AA.
XX
XX AAE10144;
XX
XX 29-NOV-2001 (first entry)
XX
XX Streptomyces noursei nystatin gene, NySk (complete).
DE
XX Polyketide synthase; PKS; macrocyclic; nystatin; PKS gene cluster;
KW antifungal; antibiotic; PKS type I.
XX
XX Streptomyces noursei.
OS
XX
XX Key Location/Qualifiers
FH 41..464
FT Domain /label= KS15 domain
FT /note= "Ketosynthase (KS) domain"
FT 578..889
FT Domain /label= ART15 domain
FT /note= "Acyltransferase (AT) domain"
FT 903..1102
FT Domain /label= DH15 domain
FT /note= "Dehydratase (DH) domain"

```

```

FT Domain 1446..1731
FT /label= ER15 domain
FT /note= "Enoylreductase (ER) domain"
FT 1740..1988
FT Domain /label= KR15 domain
FT /note= "Ketoreductase (KR) domain"
FT 2023..2096
FT Domain /label= ACP15 domain
FT /note= "Acyl carrier protein (ACP) domain"
FT 2117..2538
FT Domain /label= KS16 domain
FT /note= "Ketosynthase (KS) domain"
FT 2635..2953
FT Domain /label= ART16 domain
FT /note= "Acyltransferase (AT) domain"
FT 2967..3167
FT Domain /label= DH16 domain
FT /note= "Dehydratase (DH) domain"
FT 3257..3500
FT Domain /label= KR16 domain
FT /note= "Ketoreductase (KR) domain"
FT 3539..3612
FT Domain /label= ACP16 domain
FT /note= "Acyl carrier protein (ACP) domain"
FT 3634..4057
FT Domain /label= KS17 domain
FT /note= "Ketosynthase (KS) domain"
FT 4153..4472
FT Domain /label= ART17 domain
FT /note= "Acyltransferase (AT) domain"
FT 4486..4725
FT Domain /label= DH17 domain
FT /note= "Dehydratase (DH) domain"
FT 4997..5245
FT Domain /label= KR17 domain
FT /note= "Ketoreductase (KR) domain"
FT 5277..5350
FT Domain /label= ACP17 domain
FT /note= "Acyl carrier protein (ACP) domain"
XX
XX WO200159126-A2.
XX
XX 16-AUG-2001.
XX
XX PD
XX
XX 08-FEB-2001; 2001WO-GB00509.
XX
XX PF
XX 08-FEB-2000; 2000GB-0002840.
XX
XX PR 10-APR-2000; 2000GB-0008786.
XX
XX PR 14-APR-2000; 2000GB-0009387.
XX
XX
XX (UNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
XX (SNTF) SINTER STIFTELSEN IND TEK FORSK.
XX (ALPH-) ALPHARMA AS.
XX (SINV-) SINVENT AS.
XX (DZIE/) DZIEGLEWSKA H.
XX (ZOTC/) ZOTCHEV S B.
XX (SEKU/) SEKUROVA O N.
XX (FJAE/) FJAEVIRK E.
XX (BRAU/) BRAUTASET T.
XX (STRO/) STROM A R.
XX
XX
XX Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;
PI Valia S, Ellingsen TE, Sletta H, Gulliksen O;
XX
XX WPI: 2001-557614/62.
XX N-PSDB; AADI7186.
DR
DR N-PSDB; AADI7186.
XX
XX New nystatin polyketide synthase polynucleotides and polypeptides,
PT useful as antibiotics and antifungals -
XX
XX
XX Claim 15; Page 255-260; 266pp; English.
PS
PS The present invention relates to the cloning and sequencing of the gene
XX

```


Oy	62	DTSNPDQIQOQLNTPTPPHPTTTVHTTNGILDDATLTUTPTPOLNNVLRKAKHSHLHQ	122
Db	4244	DIGBRVAVALLDQVPAQHPLKAVHTHGVDDGLTSLTERPKMAVLHAKKFGAAHLHD	4300
Oy	122	LTQHTPLTAPVLVYSSAAATFGAPGQANYAAANAVDALAHHRHHTHLPATSIAMGTW	178
Db	4304	LTRDAGLTFTTVFSSAAASFGSPGGNGYTAANAFDLMQHRHTQALPGRSLAWGLW	4360
RESULT 14			
AAE10142	ID	AAE10142	standard; Protein; 7068 AA.
XX	AAE10142,		
AC	AAE10142,		
XX	29-NOV-2001	(first entry)	
DT	XX		
DE	Streptomyces novisei	nystatin gene, NysI (partial).	
XX			
KW	Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;		
KW	antifungal; antibiotic; PKS type I.		
XX			
OS	Streptomyces novisei.		
XX			
Key	Location/Qualifiers		
FT	34..448		
FT	/label= KS9_domain		
FT	/note= "Ketosynthase (KS) domain"		
FT	572..890		
FT	/label= AT9_domain		
FT	/note= "Acyltransferase (AT) domain"		
FT	904..1123		
FT	/label= DH9_domain		
FT	/note= "Dehydratase (DH) domain"		
FT	1443..1686		
FT	/label= KR9_domain		
FT	/note= "Ketoreductase (KR) domain"		
FT	1720..1793		
FT	/label= ACP9_domain		
FT	/note= "Acyl carrier protein (ACP) domain"		
FT	1813..2236		
FT	/label= KS10_domain		
FT	/note= "Ketosynthase (KS) domain"		
FT	2346..2664		
FT	/label= AT10_domain		
FT	/note= "Acyltransferase (AT) domain"		
FT	2678..2890		
FT	/label= DH10_domain		
FT	/note= "Dehydratase (DH) domain"		
FT	2983..3229		
FT	/label= KR10_domain		
FT	/note= "Ketoreductase (KR) domain"		
FT	3266..3339		
FT	/label= ACP10_domain		
FT	/note= "Acyl carrier protein (ACP) domain"		
FT	3358..3780		
FT	/label= KS11_domain		
FT	/note= "Ketosynthase (KS) domain"		
FT	3898..4217		
FT	/label= AT11_domain		
FT	/note= "Acyltransferase (AT) domain"		
FT	4231..4432		
FT	/label= DH11_domain		
FT	/note= "Dehydratase (DH) domain"		
FT	4523..4770		
FT	/label= KR11_domain		
FT	/note= "Ketoreductase (KR) domain"		
FT	4806..4879		
FT	/label= ACP11_domain		
FT	/note= "Acyl carrier protein (ACP) domain"		
FT	4801..5325		
FT	/label= KS12_domain		

Query Match	Similarity	Score	DB	Length	7068
Best Local Match	50.6%	Pred. No. 1.6e-37			
Matches	90	Conservative	20	Mismatches	67
				Indels	1
				Gaps	1
1	GTTLITGATGATATLTHNLTTTQPTOHLITSRGPTPHAOHLTTLOOGKIHITTT	60			
6068	CTVITITGTCIGTGHVARRLARDGAT-NLLITSRGDAAGDRLRYALEBEGAKRVTLAA	61266			
61	CDTSNPDOQLNTIPQHPRLTVITHTAGILDDATLTNTLPTQJLNNVLRKAHSAHLH	120			
6127	CDADRDALALALAELEDDAPLCVFTHTAGVDEHVDALTPENFAVLRAKTVAAHHLH	6186			
121	QLTQHTLTTLFVLYSSAAATFGAGGQNVAAAYDALAHNRHTHNLPRATSTMGTW	178			
6187	ELTADLDLALFVLFSSTAGVLAAGGQNVAAAHLDALAEHRSHGLTALSVAMGFW	6244			

RESULT 15
 ID AAM22611 standard; Protein, 4545 AA.
 XX AAM22611;
 AC AAM22611;
 XX
 DT 02-MAR-1998 (first entry)
 XX
 DE Hybrid srmG/tylG ORF1 protein.
 XX
 KM Ty lactone synthase gene cluster; tylG gene; multifunctional protein;
 XX polyketide; ty lactone synthesis; antibiotic; tylosin; hybrid protein.
 OS Streptomyces fradiae.
 XX
 XX Key
 FT Modified-site 1 Location/Qualifiers
 FT /note= "encoded by GNG"
 FT 15..418
 FT /note= "ketosynthase domain, KS'(s), from tylG"
 FT 525..882
 FT /note= "acyltransferase domain, AT(s), from tylG"
 FT 942..1025
 FT /note= "acyl carrier protein domain, ACP(s), from tylG"
 FT 1060..1483
 FT /note= "ketosynthase domain, KS1, from srmG"
 FT 1596..1953
 FT /note= "acyltransferase domain, AT1, from srmG"
 FT 2232..2416
 FT /note= "ketoreductase domain, KR1, from srmG"
 FT 2533..2616
 FT /note= "acyl carrier protein domain, ACP1, from srmG"
 FT 2641..3064
 FT /note= "ketosynthase domain, KS2, from srmG"
 FT 3184..3520
 FT /note= "acyltransferase domain, AT2, from srmG"
 FT 3546..3727
 FT /note= "dehydratase domain, DH2, from srmG"
 FT 4083..4268
 FT /note= "ketoreductase domain, KR2, from srmG"
 FT 4374..4457
 FT /note= "acyl carrier protein domain, ACP2, from srmG"
 FT Domain
 XX
 PN EP791655-A2.
 XX
 PD 27-AUG-1997.
 XX
 PF 19-FEB-1997; 97BP-0301056.
 XX
 PR 22-FEB-1996; 96US-0012078.
 XX
 XX (ELIL) LILLY & CO ELI.
 XX
 PI Dehoff BS, Kuhatosa SA, Roestek PR, Sutton XL;
 XX WPI, 1997-418046/39.
 DR N-PSDB; AAT80415.
 XX
 PT DNA encoding Streptomyces fradiae ty lactone synthase domain - for
 XX production of tylosin-related polyketide compounds
 PS Claim 23; Pages 198-212; 220pp; English.
 XX
 CC This sequence represents a hybrid protein of the invention. This
 CC sequence was created by replacing a EcoRI-ApaI fragment of srmG ORF1 with
 CC a EcoRI-SfiI fragment from tylG ORF1. The position of the residues from
 CC each of the two genes is not given in the specification. The srmG gene
 CC (see AAT80415) was isolated from Streptomyces ambifaciens, and encodes
 CC the multi-functional proteins which direct the synthesis of the
 CC polyketide plateinoid. Plateinoid is the basic building block of the

CC macrocyclic antibiotic spiramycin. The tylG gene (see AAT80413) is the
 CC ty lactone synthase gene cluster of the invention. The tylG sequence was
 CC isolated from Streptomyces fradiae, and encodes multifunctional proteins
 CC which direct the synthesis of the polyketide ty lactone. Ty lactone is the
 CC basic building block of the antibiotic tylosin. The hybrid sequence can
 CC be used to transform S. ambifaciens lacking the srmG ORF1 sequence, or S.
 CC fradiae lacking the tylG ORF1 sequence, so that they can produce
 CC polyketides. The DNA sequence can be modified so as to alter the type of
 CC carboxylic acids incorporated, the number of carboxylic acids
 CC incorporated and/or the post-condensation reactions performed, thereby
 CC resulting in novel tylosin-related polyketides.

XX
 SQ Sequence 4545 AA;

Query Match 46.1%; Score 439; DB 18; Length 4545;

Best Local Similarity 49.4%; Pred. No. 1.6e-37;

Matches 88; Conservative 23; Mismatches 67; Indels 0; Gaps 0;

Qy 1 GTTLITGGTGAATHTLTHLTTHPTQTHLLTSRTGPHRPHAOHTTQLOQKGIHLTTT 60

Db 4080 GTVLITGGTGAATHTLTHLTTHPTQTHLLTSRTGPHRPHAOHTTQLOQKGIHLTTT 4139

Qy 61 CDTSNPDLOQLNTIPPOHPLTTVHTAGILDATLTNLTPTQANNVLRKXSHLH 120

Db 4140 CDLTDEAVKAKAAVAPSDHPLTGVTGVDGALTGLTRQRLDTLVRPKDAVRNLH 4199

Qy 121 QLTQHPPLTFVLYSSAATFGAPGOANYAANAVALDAAHHTHTHLPTSIAWGTM 178

Db 4200 EATLDRLRAFLVFSAAAGLGRPGQSTAAANAVALDAAHHTHTHLPTSIAWGTM 4257

Search completed: June 17, 2003, 13:02:30
 Job time : 9.71975 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:53:02 ; Search time 14.1889 Seconds
(without alignments)
4458.180 Million cell updates/sec

Title: US-09-914-286-3_COPY_1050_1356

Perfect score: 1598
Sequence: 1 VFVFGGQMGWAGVRLLA.....ABDVTAIGSLRGDNTRRF 307

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing filter 45 summaries

Database :

1: SPTRMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1598	100.0	3972	2 Q9S0R8	Q9S0R8 streptomyc
2	1598	100.0	5532	2 Q9S0R4	Q9S0R4 streptomyc
3	1366	85.5	6239	2 Q9S0R7	Q9S0R7 streptomyc
4	1227.5	76.8	3626	2 Q9F7T9	Q9F7T9 streptomyc
5	945	59.1	3970	2 Q93HT8	Q93HT8 streptomyc
6	891.5	55.8	3816	2 Q93HT3	Q93HT3 streptomyc
7	867	54.3	4685	2 Q93HJ2	Q93HJ2 streptomyc
8	867	54.3	9510	2 Q93HX9	Q93HX9 streptomyc
9	866	54.2	3613	2 Q93HJ1	Q93HJ1 streptomyc
10	861.5	53.9	4150	2 Q93HT4	Q93HT4 streptomyc
11	858	53.7	3190	2 Q93NM7	Q93NM7 streptomyc
12	852	53.3	2595	2 Q9ALM6	Q9ALM6 saccharopol
13	848.5	53.1	3170	2 Q9ALM4	Q9ALM4 saccharopol
14	848.5	53.1	5060	2 Q52545	Q52545 amycolatops
15	848.5	53.1	5069	2 Q52789	Q52789 amycolatops
16	844.5	52.8	1041	2 Q93C90	Q93C90 actinomadura

17	841.5	52.7	9507	2 Q9EW41	Q9EW41 streptomyc
18	841	52.6	1907	2 Q93H19	Q93H19 streptomyc
19	837.5	52.4	8817	2 Q53840	Q53840 polyangium
20	835.5	52.3	3654	2 Q30766	Q30766 streptomyc
21	832	52.1	3546	2 Q9F830	Q9F830 microonosp
22	832	52.1	4881	2 Q9S0R3	Q9S0R3 streptomyc
23	828.5	51.8	3192	2 Q914W4	Q914W4 streptomyc
24	827	51.8	3562	2 Q9F829	Q9F829 microonosp
25	826	51.7	9477	2 Q914X3	Q914X3 streptomyc
26	816	51.1	3524	2 Q93H86	Q93H86 streptomyc
27	815.5	51.0	4735	2 Q54666	Q54666 amycolatops
28	813	50.9	4840	2 Q93HJ4	Q93HJ4 streptomyc
29	812.5	50.8	1762	2 Q52546	Q52546 amycolatops
30	812.5	50.8	1763	2 Q52790	Q52790 amycolatops
31	809.5	50.7	3201	2 Q9F828	Q9F828 microonosp
32	807	50.5	7576	2 Q9ZG44	Q9ZG44 streptomyc
33	806.5	50.5	3413	2 Q54593	Q54593 amycolatops
34	806.5	50.5	4290	2 Q9WXC0	Q9WXC0 microonosp
35	806	50.4	6146	2 Q93HJ5	Q93HJ5 streptomyc
36	805	50.4	1728	2 Q54591	Q54591 amycolatops
37	801.5	50.2	4498	2 Q93H85	Q93H85 streptomyc
38	800	50.1	3939	2 Q93HJ3	Q93HJ3 streptomyc
39	795	49.7	7525	2 Q9KTE0	Q9KTE0 streptomyc
40	790.5	49.5	2126	16 P94996	P94996 mycobacteri
41	786.5	49.2	6315	2 Q9ADL6	Q9ADL6 polyangium
42	785	49.1	3352	2 Q93H83	Q93H83 streptomyc
43	784	49.1	4151	16 Q53490	Q53490 mycobacteri
44	775.5	48.5	1937	2 Q30482	Q30482 streptomyc
45	774	48.4	5588	2 Q9ALM2	Q9ALM2 saccharopol

ALIGNMENTS

RESULT 1

ID Q9S0R8 PRELIMINARY; PRT; 3972 AA.

AC Q9S0R8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Type I polyketide synthase AVES 1.
GN AVEAL.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycesetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9380546; PubMed=10449723;
RT Ireda H., Nonomiy T., Usami M., Ohta T., Omura S.,
RT "Organization of the biosynthetic gene cluster for the polyketide
RT anthelmintic macroide avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL; AB032367; BAA84474.1; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantane-attach.
DR InterPro; IPR000834; Zn carboxypept.
DR Pfam; PF00698; Acyl_transf. 3.
DR Pfam; PF0109; ketoacyl-synt. 2.
DR Pfam; PF02801; ketoacyl-synt. C; 2.
DR Pfam; PF0550; pp-binding; 3.
DR PROSITE; PS50075; ACP_DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL SYNTHASE; 2.
DR PROSITE; PS00133; CARBOXYPEPT ZN 2; UNKNOWN_2.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 3972 AA; 416852 MW; 2A2936958032B1C3 CRC64;

Query Match 100.0%; Score 1598; DB 2; Length 3972;
Best Local Similarity 100.0%; Pred. No. 2, 2e-105;
Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 VFVPPGGGGWAGVLLASSPVFARRMQACEBALAPWVMSVVDILRRDAGDAWVERA 60
DB 1050 VFVPPGGGGWAGVLLASSPVFARRMQACEBALAPWVMSVVDILRRDAGDAWVERA 1109
QY 61 DVVQPVLFPSVWVSLAALMRSYGIPEPDVVLGHSQGEIAAAHVCGALSILKDAKTVALLRSRA 120
DB 1110 DVVQPVLFPSVWVSLAALMRSYGIPEPDVVLGHSQGEIAAAHVCGALSILKDAKTVALLRSRA 1169
QY 121 LAARGRGGAASVPLPAOEVEQLIGERMAGRLWVAANGPSTAVSGDAEAVDEVLAYCA 180
DB 1170 LAARGRGGAASVPLPAOEVEQLIGERMAGRLWVAANGPSTAVSGDAEAVDEVLAYCA 1229
QY 181 GTGVARARIPVDYASHCPHVQPLREBELLELGDISPQSGVPFFSTVEGTWLDTTTLDDAA 240
DB 1230 GTGVARARIPVDYASHCPHVQPLREBELLELGDISPQSGVPFFSTVEGTWLDTTTLDDAA 1289
QY 241 YWYRNLIHQPVRFSDAVQALADDGHRVFEVSPHPTLVPAIEDTTEPTAEVTAIGSLIRG 300
DB 1290 YWYRNLIHQPVRFSDAVQALADDGHRVFEVSPHPTLVPAIEDTTEPTAEVTAIGSLIRG 1349
QY 301 DNDTRRF 307
DB 1350 DNDTRRF 1356

RESULT 2
09S0R4 PRELIMINARY; PRT; 5532 AA.
AC 09S0R4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 3.
GN AVEA3.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99380548; PubMed=10449723;
RA Ikeda H., Nonomiya T., Usami M., Ohca T., Omura S.;
RT "Organization of the biosynthetic gene cluster for the polyketide
RT anthelmintic macrocyclic avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL: AB032367; BAA84478.1;
DR InterPro: IPR001227; Ac_transferase.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR003880; Pantane attach.
DR InterPro: IPR000834; Zn_carboxypept.
DR Pfam: PF00698; Acyl_transf. 3.
DR Pfam: PF02801; ketoacyl-synt. 3.
DR Pfam: PF02801; ketoacyl-synt. 3.
DR Pfam: PF05075; pp-binding. 3.
DR PROSITE: PS50075; ACP DOMAIN. 3.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 3.
DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
DR PROSITE: PS00012; PHOSPHOPANTHETINE; 3.
KW Phosphopantetheine; transferase.
SQ SEQUENCE 5532 AA; 575193 MW; EC7B2069D44A4057 CRC64;

Query Match 100.0%; Score 1598; DB 2; Length 5532;
Best Local Similarity 100.0%; Pred. No. 3.3e-105;
Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 121 LAARGRGGAASVPLPAOEVEQLIGERMAGRLWVAANGPSTAVSGDAEAVDEVLAYCA 180
DB 723 LAARGRGGAASVPLPAOEVEQLIGERMAGRLWVAANGPSTAVSGDAEAVDEVLAYCA 782
QY 181 GTGVARARIPVDYASHCPHVQPLREBELLELGDISPQSGVPFFSTVEGTWLDTTTLDDAA 240
DB 783 GTGVARARIPVDYASHCPHVQPLREBELLELGDISPQSGVPFFSTVEGTWLDTTTLDDAA 842
QY 241 YWYRNLIHQPVRFSDAVQALADDGHRVFEVSPHPTLVPAIEDTTEPTAEVTAIGSLIRG 300
DB 843 YWYRNLIHQPVRFSDAVQALADDGHRVFEVSPHPTLVPAIEDTTEPTAEVTAIGSLIRG 902
QY 301 DNDTRRF 307
DB 903 DNDTRRF 909

RESULT 3
09S0R7 PRELIMINARY; PRT; 6239 AA.
AC 09S0R7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Type I polyketide synthase AVES 2.
GN AVEA2.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99380548; PubMed=10449723;
RA Ikeda H., Nonomiya T., Usami M., Ohca T., Omura S.;
RT "Organization of the biosynthetic gene cluster for the polyketide
RT anthelmintic macrocyclic avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL: AB032367; BAA84475.1;
DR InterPro: IPR001227; Ac_transferase.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR003880; Pantane attach.
DR InterPro: IPR000169; SHprot_acsite.
DR InterPro: IPR000834; Zn_carboxypept.
DR Pfam: PF00698; Acyl_transf. 4.
DR Pfam: PF00109; ketoacyl-synt. 4.
DR Pfam: PF02801; ketoacyl-synt. 4.
DR Pfam: PF05075; pp-binding. 4.
DR PROSITE: PS50075; ACP DOMAIN. 4.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 4.
DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_4.
DR PROSITE: PS00012; PHOSPHOPANTHETINE; 3.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
KW Phosphopantetheine; transferase.
SQ SEQUENCE 6239 AA; 666283 MW; A706280940B502FA CRC64;

Query Match 85.5%; Score 1366; DB 2; Length 6239;
Best Local Similarity 81.8%; Pred. No. 1.6e-88;
Matches 251; Conservative 25; Mismatches 31; Indels 0; Gaps 0;

```

ID	Q93H18	PRELIMINARY;	PRT	3970 AA.
Db	4995 DTGARAKRIIPVDYASHCHVPRLHMLHLGLDITRPOSTVYFSTVSGTLLDITTLDDAA	5054		
Qy	241 YWYRNLLHQPVRPSDAVQALADGHHVPEVSPHPTLVPAIEDTTEDTAEDVTAIGSLRG	300		
Db	5055 YWYRNLLHQPVRPSDAVQALADGHHVPEVSPHPTLVPAIEDTTEDTAEDVTAIGSLRG	5114		
Qy	301 DNDTRRF 307			
Db	5115 DNDTRRF 5121			
RESULT 4				
Q9F7T9	PRELIMINARY;	PRT;	3626 AA.	
ID	Q9F7T9	AC	Q9F7T9;	
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Avermectin polyketide synthase (Fragment).			
OC	Streptomyces avermectilis.			
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
OC	Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.			
OX	NCBI_Taxid=33903;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC12711.			
RA	Hong Y.-S., Lee J.J.;			
RT	"Targeted Gene Disruption of the avermectin O-methyltransferase gene			
RT	and polyketide synthase gene from Streptomyces avermectilis."			
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.			
RL	EMBL; AF215943; AAG09812.1; -			
DR	InterPro; IPR001277; AC_transferase.			
DR	InterPro; IPR00794; Ketoacyl-synt.			
DR	InterPro; IPR003880; Pantone attach.			
DR	Pfam; PF00698; ACYL_transf. 3.			
DR	Pfam; PF00109; ketoacyl-synt. 2.			
DR	Pfam; PF02801; ketoacyl-synt_C1 2.			
DR	Pfam; PF00550; DP-binding; 2.			
DR	PROSITE; PS50075; ACP DOMAIN; 2.			
DR	PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.			
DR	PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.			
FT	NON_TER			
SO	SEQUENCE 3626 AA; 380557 MW; 6272F5F08C1A8D0 CRC64;			
Query Match				
Best Local Similarity 76.8%; Score 1227.5; DB 2; Length 3626;				
Matches 245; Conservative 5; Mismatches 17; Indels 21; Gaps 3;				
Qy	1 VFVPPGGGGVAGMGVRLILASSPVFARRMQACEALAPWDMVSVDILRRDAGDAVWERA	60		
Db	1037 VFVFPWGGGOWAGMGVRLILCLTRVARRMQACEALAPWDMVSVDILRRDAGDAVWEOA	10366		
Qy	61 DVNCPVLFSSVWVSLAALWRSYGIEPDALVGHSGGIIAAHYCGSLSLDAKTAVALRSRA	120		
Db	1097 DVNCPVLFSSVWVSLAALWRSYGIEPDALVGHSGGIIAAHYCGSLSLDAKTAVALRSRA	11511		
Qy	121 LAAVAGRGMAVPLPAAQVEVOLIERWAGRLVNAVNGPRSTAVSGDAEAVDEVLAICYA	180		
Db	1152 -----LPQVEVOLIERWAGRLVNAVNGPRSTAVSGDAEAVDEVLAICYA	11968		
Qy	181 GTGVARRIPVDYASHCHVPRLRELLLELGLDISPOPQVGFSTVEGTWLDITTLDDAA	240		
Db	1197 GTGVARRIPVDYASHCHVPRLRELLLELGLDISPOPQVGFSTVEGTWLDITTLDDAA	12555		
Qy	241 YWYRNLLHQPVRPSDAVQALADGHHVPEVSPHPTLVPAIEDTTEDTA 288			
Db	1256 YWYRNLLHQPVRPSDAVQALADGHHVPEVSPHPTLVPAIEDTTEDTA 1303			

[illegible]

DT 01-JUN-2002 (TREMBLER. 21, Last annotation update)
 DE 8, 9a-deoxyoleandolide synthase 2.
 GN OLEA11.
 OS Streptomyces antibioticus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20363406; PubMed=10908114;
 RA Shah S., Xue Q., Tang L., Carney J.R., Betlach M., McDaniel R.;
 RT "Cloning, Characterization, and Heterologous Expression of a
 RT Polypeptide Synthase and P-450 Oxidase Involved in the Biosynthesis of
 RT the Antibiotic Oleandomycin."
 RL J. Antibiot. 53:502-508(2000).
 DR EMBL; AF220951; AAF82409.1; -
 DR HSP; P25715; IMLA.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh_zn_family.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantone attach.
 DR InterPro; IPR002364; OOR_zeta_crystal.
 DR Pfam; PF00698; Acyl_transf; 2.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF02801; ketoacyl-synt_C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHEINE; UNKNOWN_2.
 DR PROSITE; PS01162; OOR_ZETA_CRYSTAL; 1.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 3816 AA; 400235 MW; 6741BFD35F9B5C72 CRC64;

Query Match 55.8%; Score 891.5; DB 2; Length 3816;
 Best Local Similarity 58.6%; Pred. No. 8e-55;
 Matches 180; Conservative 31; Mismatches 93; Indels 3; Gaps 2;

QY 1 VFVPPGGGQWAGVRLASSPVFARMQACEALAPWVMSVVDILRDAGDAWVERA 60
 DB 567 VFVPPGGGQWAGVRLASSPVFARMQACEALAPWVMSVVDILRDAGDAWVERA 624
 QY 61 DVVQPVLFVSVMSLAALMRSYGIIEPDVAVLGHSGEIAAAHVCGALSLDKAAKTVALRSR 120
 DB 625 DVVQPVLFVSVMSLAALMRSYGIIEPDVAVLGHSGEIAAAHVCGALSLDKAAKTVALRSR 684
 QY 121 LAAVRGSGMAVSLPDAOEVEQLIGERMAQLMVAANVPSTAVSGDAEAVDEVLAAYCA 180
 DB 685 IARIAGGGGMAVSLPDAOEVEQLIGERMAQLMVAANVPSTAVSGDAEAVDEVLAAYCA 743
 QY 181 GTGVARRRIPVDYASHCHVQPLREBELLELDGISPOPSGVFPFSTVEGTWLDTTTDA 240
 DB 744 REGRARARVPDYASHCHVQPLREBELLELDGISPOPSGVFPFSTVEGTWLDTTTDA 803
 QY 241 YVYENLHQPVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDAEDVTAGSLRG 300
 DB 804 YVYENLHQPVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDAEDVTAGSLRG 863
 QY 301 DNDTRRF 307
 DB 864 EGGIGRF 870

RESULT 7
 Q93HJ2 PRELIMINARY; PRT; 4685 AA.
 AC Q93HJ2;
 DT 01-DEC-2001 (TREMBLER. 19, Created)
 DT 01-DEC-2001 (TREMBLER. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLER. 21, Last annotation update)
 DE Modular polypeptide synthase.
 GN OLEA6.
 OS Streptomyces avermitilis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinoue M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: Deducing the ability of producing secondary
 RT metabolites."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL; AB070940; BAB69195.1; -
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Pantone attach.
 DR Pfam; PF00106; adh_short; 3.
 DR Pfam; PF00109; ketoacyl-synt; 3.
 DR Pfam; PF02801; ketoacyl-synt_C; 3.
 DR Pfam; PF00550; pp-binding; 3.
 DR PROSITE; PS50075; ACP DOMAIN; 3.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_3.
 DR PROSITE; PS00012; PHOSPHOPANTHEINE; UNKNOWN_3.
 KM Phosphopantetheine.
 SQ SEQUENCE 4685 AA; 489355 MW; 2FA34215A39A0DB8 CRC64;

Query Match 54.3%; Score 867; DB 2; Length 4685;
 Best Local Similarity 58.0%; Pred. No. 5.9e-53;
 Matches 178; Conservative 37; Mismatches 90; Indels 2; Gaps 2;

QY 1 VFVPPGGGQWAGVRLASSPVFARMQACEALAPWVMSVVDILRDAGDAWVERA 60
 DB 2096 VFVPPGGGQWAGVRLASSPVFARMQACEALAPWVMSVVDILRDAGDAWVERA 2155
 QY 61 DVVQPVLFVSVMSLAALMRSYGIIEPDVAVLGHSGEIAAAHVCGALSLDKAAKTVALRSR 119
 DB 2156 DVVQPVLFVSVMSLAALMRSYGIIEPDVAVLGHSGEIAAAHVCGALSLDKAAKTVALRSR 2215
 QY 120 ALAAVGRSGMAVSLPDAOEVEQLIGERMAQLMVAANVPSTAVSGDAEAVDEVLAAYC 179
 DB 2216 VLKELSGKGMVSLPDAOEVEQLIGERMAQLMVAANVPSTAVSGDAEAVDEVLAAYC 2274
 QY 180 AGTGVARRIPVDYASHCHVQPLREBELLELDGISPOPSGVFPFSTVEGTWLDTTTDA 239
 DB 2275 EGGVARRIPVDYASHCHVQPLREBELLELDGISPOPSGVFPFSTVEGTWLDTTTDA 2334
 QY 240 AYVYENLHQPVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDAEDVTAGSLRG 299
 DB 2335 RYVYENLHQPVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDAEDVTAGSLRG 2394
 QY 300 GNDTRRF 306
 DB 2395 DDGADR 2401

RESULT 8
 Q93NX9 PRELIMINARY; PRT; 9510 AA.
 AC Q93NX9;
 DT 01-DEC-2001 (TREMBLER. 19, Created)
 DT 01-DEC-2001 (TREMBLER. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLER. 21, Last annotation update)
 DE AmpH1.
 GN AMPH1.
 OS Streptomyces nodosus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=40318;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Caffrey P., Lynch S.V., Flood E.M., Finnian S.M., O'Leary M.;
 RT "The amphotericin biosynthetic gene cluster from Streptomyces
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AF557202; AAK73501.1; -
 DR InterPro: IPR001227; Acetyltransferase.
 DR InterPro: IPR002198; ADH short.
 DR InterPro: IPR004410; FAD.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR001005; Myb_DNA_binding.
 DR InterPro: IPR003880; Pantane_attach.
 DR Pfam: PF00698; Acyl_transfer; 6.
 DR Pfam: PF00109; ketoacyl-synt; 1.
 DR Pfam: PF02801; ketoacyl-synt C; 6.
 DR Pfam: PF00550; pp-binding; 6.
 DR TIGRfam: TIGR00128; ladd; 6.
 DR PROSITE: PS50075; ACP_DOMAIN; 6.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_6.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_6.
 KM Phosphopantetheine.
 SQ SEQUENCE 9510 AA; 992817 MW; 1A8E003A15D478CF CRC64;

Query Match 54.3%; Score 867; DB 2; Length 9510;
 Best Local Similarity 55.5%; Pred. No. 1.4e-53;

Matches 171; Conservative 48; Mismatches 87; Indels 2; Gaps 2;

QY 1 VFVPPGCGGQWAGNGVRLASSPVFARMQACEBALAPWDMVSVDILRRDAGDAVERA 60
 DB 3851 VFVPPGCGGQWAGNGVRLASSPVFARMQACEBALAPWDMVSVDILRRDAGDAVERA 3910
 QY 61 DVVQPVLFVSVVSLAALMRSYGIEPDVVLGHSGGEIAAAHVCGALSLKDAKTVALSRA 120
 DB 3911 DVVQPVLFVSVVSLAALMRSYGIEPDVVLGHSGGEIAAAHVCGALSLKDAKTVALSRA 3970
 QY 121 LA-AVRGRGGMASVPLPAQVEVQLIGRMAGRLVAAVNGPRSTAVSGDAEAVDEVILAYC 179
 DB 3971 IGRSLAGRGGMASVPLPAQVEVQLIGRMAGRLVAAVNGPRSTAVSGDAEAVDEVILAYC 4029
 QY 180 AGTGVARRRIPVUYASHCHPVQPLREBELLELDISPOSGVPPFSTVEGTWIDTTTLDA 239
 DB 4030 TAEIRARRRIPVUYASHCHPVQPLREBELLELDISPOSGVPPFSTVEGTWIDTTTLDA 4089
 QY 240 AYWYRNLRHQPRESDAVQALADGHRVFEVSPHPTLPVPIEDTTEPTAEDVTALISGLRR 299
 DB 4090 GYWRYSLRGRVLPADAVRDLIAADHRAFIIEVSSHPLAMSVODMIDAGVAGVAGTGLRR 4149
 QY 300 GDNDRRF 307
 DB 4150 DNGGLDRF 4157

RESULT 9

Q93HU1 PRELIMINARY; PRT; 3613 AA.

AC Q93HU1 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Modular polyketide synthase.
 GN OMA7.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinoe M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,

RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: Deducting the ability of producing secondary
 RT metabolites";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL, AB070940; BAB69196.1; -
 DR InterPro: IPR001227; Acetyltransferase.
 DR InterPro: IPR002198; ADH short.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Pantane_attach.
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00698; Acyl_transfer; 2.
 DR Pfam: PF00109; ketoacyl-synt; 2.
 DR Pfam: PF02801; ketoacyl-synt C; 2.
 DR Pfam: PF00550; pp-binding; 2.
 DR Pfam: PF00975; Thioesterase; 1.
 DR PROSITE: PS50075; ACP_DOMAIN; 2.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_2.
 KM Phosphopantetheine.
 SQ SEQUENCE 3613 AA; 379236 MW; 6FA85CAB1508442 CRC64;

Query Match 54.2%; Score 866; DB 2; Length 3613;
 Best Local Similarity 56.8%; Pred. No. 5e-53;

Matches 175; Conservative 46; Mismatches 85; Indels 2; Gaps 2;

QY 1 VFVPPGCGGQWAGNGVRLASSPVFARMQACEBALAPWDMVSVDILRRDAGDAVERA 60
 DB 573 VFVPPGCGGQWAGNGVRLASSPVFARMQACEBALAPWDMVSVDILRRDAGDAVERA 632
 QY 61 DVVQPVLFVSVVSLAALMRSYGIEPDVVLGHSGGEIAAAHVCGALSLKDAKTVALSRA 119
 DB 633 DVVQPVLFVSVVSLAALMRSYGIEPDVVLGHSGGEIAAAHVCGALSLKDAKTVALSRA 692
 QY 120 ALAAVRGGMASVPLPAQVEVQLIGRMAGRLVAAVNGPRSTAVSGDAEAVDEVILAYC 179
 DB 693 VLEKLSGKGGMASVPLPAQVEVQLIGRMAGRLVAAVNGPRSTAVSGDAEAVDEVILAYC 751
 QY 180 AGTGVARRRIPVUYASHCHPVQPLREBELLELDISPOSGVPPFSTVEGTWIDTTTLDA 239
 DB 752 EESGVARRRIPVUYASHCHPVQPLREBELLELDISPOSGVPPFSTVEGTWIDTTTLDA 811
 QY 240 AYWYRNLRHQPRESDAVQALADGHRVFEVSPHPTLPVPIEDTTEPTAEDVTALISGLRR 299
 DB 812 RYWTNLRERVFEVSPHPTLPVPIEDTTEPTAEDVTALISGLRR 871
 QY 300 GDNDRRF 307
 DB 872 DEGGADR 879

RESULT 10

Q9KIV4 PRELIMINARY; PRT; 4150 AA.

AC Q9KIV4 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 8,8a-deoxyoleandolid synthase 1.
 GN OLEA1.
 OS Streptomyces antibioticus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1890;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20363406; PubMed=10908114;
 RA Shah S., Xue Q., Tang L., Carney J.R., Beclach M., McDaniel R.;
 RT "Cloning, Characterization, and Heterologous Expression of a
 RT Polyketide Synthase and P-450 Oxidase Involved in the Biosynthesis of
 RT the Antibiotic Oleandomycin";
 RL J. Antibiot. 53:502-508(2000).

DR EMBL; AF220951; AAF82408.1; -
 DR HSSP; P25715; IMLA.
 DR InterPro; IPR002106; AALRNA_LigaseII.
 DR InterPro; IPR001227; Ac_transferase.
 DR InterPro; IPR004410; fadD.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Pantane attach.
 DR InterPro; IPR000169; Shprt acsIc.
 DR Pfam; PF00698; Acyl_transf_3.
 DR Pfam; PF00109; ketoacyl-synt_3.
 DR Pfam; PF02801; ketoacyl-synt_C_3.
 DR Pfam; PF00550; pp-binding; 3.
 DR TIGRFAMs; TIGR00128; fadD; 3.
 DR PROSITE; PS00339; AA TRNA LIGASE_II_2; UNKNOWN_1.
 DR PROSITE; PS50075; ACP DOMAIN; 3.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 3.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
 DR Phosphatetheine; Transferase.
 SQ SEQUENCE 4150 AA; 435261 MW; 9383296C4C1647D CRC64;

Query Match 53.9%; Score 861.5; DB 2; Length 4150;
 Best Local Similarity 56.7%; Pred. No. 1.2e-52;
 Matches 174; Conservative 35; Mismatches 95; Indels 3; Gaps 2;

QY 1 VFVPPGGGQWAGNGVRLASSPVFARMQACEBALAPWVDSVVDILRRDAGDAWERA 60
 DB 3076 VFVPPGGGQWAGNGVRLASSPVFARMQACEBALAPWVDSVVDILRRDAGDAWERA 60
 QY 61 DVVQPVLFVSVVSLAALMRSYGIEPDVAVLGHSGEIAAAHVCGALSLKDAKTVALLRSRA 120
 DB 3134 DVVQPVLFVSVVSLAALMRSYGIEPDVAVLGHSGEIAAAHVCGALSLKDAKTVALLRSRA 120
 QY 121 LAARGGGMASVPLPAQVEEQLIGERWAGRLMVAANGPRSTAVSGDAEAVDEVLYAC 180
 DB 3194 IARIAGGGGAVSVGLSARVETML-DYTGGRVSAVAANGPRSTAVSGDAEAVDEVLYAC 180
 QY 181 GTGVARRIRPVVYASHCPVQPLREBELLELDISPOPSGVFPFSTEGVLTDTTLLDAA 240
 DB 3253 REGVARRIRPVVYASHCPVQPLREBELLELDISPOPSGVFPFSTEGVLTDTTLLDAA 240
 QY 3194 IARIAGGGGAVSVGLSARVETML-DYTGGRVSAVAANGPRSTAVSGDAEAVDEVLYAC 3252
 DB 241 YWYNLHQPVRFSDAVQALADGHRVPEVSPHPTLVPAIEDTETAEDVTAIGSLRG 300
 QY 3133 YMFNLRRETVFQAVLEGVLAQGMGAFAVECSPHPLVVGITETLDTFDADAVALSLRRD 3372
 DB 301 DNDTRRF 307
 DB 3373 EGGIDRF 3379

RESULT 11

Q93NW7 PRELIMINARY; PRT; 3190 AA.
 ID Q93NW7
 AC Q93NW7
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE AmphB.
 GN AMPHB.
 OS Streptomyces nodosus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=40318;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Catfrey P., Lynch S.V., Flood E.M., Finnan S.M., O'Lyryk M.,
 RT "The amphotericin biosynthetic gene cluster from Streptomyces
 RT nodosus".
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF357202; AAK73513.1;
 DR InterPro; IPR001227; Ac_transferase.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR000794; ketoacyl-synt.

DR InterPro; IPR003880; Pantane attach.
 DR Pfam; PF00698; Acyl_transf_2.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF02801; ketoacyl-synt_C_3; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_2.
 DR Phosphatetheine.
 SQ SEQUENCE 3190 AA; 334060 MW; EC005DABF792D898 CRC64;

Query Match 53.7%; Score 858; DB 2; Length 3190;
 Best Local Similarity 55.2%; Pred. No. 1.6e-52;
 Matches 170; Conservative 44; Mismatches 92; Indels 2; Gaps 2;

QY 1 VFVPPGGGQWAGNGVRLASSPVFARMQACEBALAPWVDSVVDILRRDAGDAWERA 60
 DB 573 VFVPPGGGQWAGNGVRLASSPVFARMQACEBALAPWVDSVVDILRRDAGDAWERA 60
 QY 61 DVVQPVLFVSVVSLAALMRSYGIEPDVAVLGHSGEIAAAHVCGALSLKDAKTVALLRSRA 120
 DB 633 DVVQPVLFVSVVSLAALMRSYGIEPDVAVLGHSGEIAAAHVCGALSLKDAKTVALLRSRA 120
 QY 121 LA-ARGGGMASVPLPAQVEEQLIGERWAGRLMVAANGPRSTAVSGDAEAVDEVLYAC 179
 DB 693 IGRITAGRGGMASVPLPAQVEEQLIGERWAGRLMVAANGPRSTAVSGDAEAVDEVLYAC 179
 QY 180 AGTVARRIRPVVYASHCPVQPLREBELLELDISPOPSGVFPFSTEGVLTDTTLLDAA 239
 DB 752 TAGGVARRIRPVVYASHCPVQPLREBELLELDISPOPSGVFPFSTEGVLTDTTLLDAA 239
 QY 240 AYWYNLHQPVRFSDAVQALADGHRVPEVSPHPTLVPAIEDTETAEDVTAIGSLRG 299
 DB 812 GYVFRSLRGVLPADHRIGLDTERHAFLEVSHPVLAANSQMDIDAGVGAAGTLRR 871
 QY 300 GDNDRRF 307
 DB 872 DNGIDRF 879

RESULT 12

Q9ALM6 PRELIMINARY; PRT; 2595 AA.
 ID Q9ALM6
 AC Q9ALM6
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Polypeptide synthase loading and extender module 1.
 GN SPNA.
 OS Saccharopolyspora spinosa.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;
 OX NCBI_TaxID=60894;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21257765; PubMed=11358695;
 RA Waldron C., Matsushima P., Rostek P.R. Jr., Broughton M.C.,
 RA Turner J., Madduri K., Crawford K.P., Merlo D.J., Baltz R.H.,
 RT "Cloning and analysis of the spinosa biosynthetic cluster of
 RT Saccharopolyspora spinosa".
 RL Chem. Biol. 8:487-499(2001).
 DR EMBL; AY007564; AAG23264.1;
 DR InterPro; IPR001227; Ac_transferase.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Pantane attach.
 DR Pfam; PF00698; Acyl_transf_2.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00109; ketoacyl-synt; 2.
 DR Pfam; PF02801; ketoacyl-synt_C_3; 2.
 DR Pfam; PF00550; pp-binding; 2.

DR PROSITE; PSS0075; ACP DOMAIN; 2.
 DR PROSITE; PSS0066; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE; PSS0012; PHOSPHOPANTHETINE; UNKNOWN_1.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 2595 AA; 273663 MW; 52EF0BBA4D1FA62 CRC64;

Query Match 53.3%; Score 852; DB 2; Length 2595;
 Best Local Similarity 55.2%; Pred. No. 3.3e-52;
 Matches 169; Conservative 47; Mismatches 88; Indels 2; Gaps 2;

QY 1 VFVPPGGGQWAGMAGVRLASSPVFARRMQACEBALPWYDMSVVDLRRDAGAWERA 60
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 527 VFVPPGGGQWIGRQLLETSEVFAASMSDCADALPHLDWSLLDVLRNAAGAAHLDDH 586
 QY 61 DVVQPVLFSSVWVSLAALMRSGYIEPDVGLHSGGEIAAAHVCALSLKDAKTAALRSRA 120
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 587 DVVQPVLFALVWVSLAELMRSGVPRVAVVGHSGEIAACVAGLSVRDAAAVVAVSRL 646
 QY 121 LAAVRGSGMASVPLPAQVEYQILGERWAGRLVVAANGPRSTAVSGDAEAVDEVLAYCA 180
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 647 LTAAGSGAASLQHPAEVYQIL-LPWRDRIGVAGVNGPSSTLVSGDREAMELLAECA 705
 QY 181 GTGVARRIPVDYASHCPHVPQPLREBELLEDGISPPSGVPFSTVEGTMLDTTIDAA 240
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 706 DRELRRMRIPVEYASHSPHIEVNDLGLLAPVPRPTGSPYISTTGDLDD-RPRDAD 764
 QY 241 YWYRNLIHQVRFSDVAQALADDGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 300
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 765 YWYRNLIHQVRFSDVAQALADDGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 824
 QY 301 DNDTRR 306
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 825 EGGWRQ 830

RESULT 13

Q9ALM4 PRELIMINARY; PRT; 3170 AA.

AC Q9ALM4;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Polypeptide synthase extender modules 3-4.
 GN SPNC.
 OS Saccharopolyspora spinosa.
 OC Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae;
 OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae;
 OC Saccharopolyspora.
 OX NCBI_TaxID=60894;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2125765; PubMed=11358695;
 RA Waldron C., Matsushita P., Roetzel P.R. Jr., Broughton M.C.,
 Turner J., Madduri K., Crawford K.P., Merlo D.J., Baltz R.H.;
 RT "Cloning and analysis of the spinosad biosynthetic cluster of
 RT Saccharopolyspora spinosa."
 RL Chem. Biol. 8:487-499(2001).
 DR EMBL; AY007564; AAG3266.1; -.
 DR HSP; P25715; IMLA.
 DR Interpro; IPR001227; AC_transferase.
 DR Interpro; IPR002198; ADH_short.
 DR Interpro; IPR000794; ketoacyl-acyl.
 DR Interpro; IPR003880; Pantene_attach.
 DR Pfam; PF00698; Acyl_transfer; 2.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00109; ketoacyl-acyl; 2.
 DR Pfam; PF02801; ketoacyl-acyl; C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PSS0075; ACP DOMAIN; 2.
 DR PROSITE; PSS0066; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE; PSS0012; PHOSPHOPANTHETINE; 2.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 3170 AA; 334873 MW; 205D519FDFPA3249C CRC64;

Query Match 53.1%; Score 848.5; DB 2; Length 3170;
 Best Local Similarity 56.2%; Pred. No. 7.6e-52;
 Matches 172; Conservative 35; Mismatches 98; Indels 1; Gaps 1;

QY 1 VFVPPGGGQWAGMAGVRLASSPVFARRMQACEBALPWYDMSVVDLRRDAGAWERA 60
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 563 VFVPPGGGQWAGMAGVRLASSPVFARRMQACEBALPWYDMSVVDLRRDAGAWERA 622
 QY 61 DVVQPVLFSSVWVSLAALMRSGYIEPDVGLHSGGEIAAAHVCALSLKDAKTAALRSRA 120
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 623 DVVQPVLFSSVWVSLAALMRSGYIEPDVGLHSGGEIAAAHVCALSLKDAKTAALRSRA 120
 QY 121 LAAVRGSGMASVPLPAQVEYQILGERWAGRLVVAANGPRSTAVSGDAEAVDEVLAYCA 180
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 683 LKLSGRGSGMASVPLPAQVEYQILGERWAGRLVVAANGPRSTAVSGDAEAVDEVLAYCA 741
 QY 181 GTGVARRIPVDYASHCPHVPQPLREBELLEDGISPPSGVPFSTVEGTMLDTTIDAA 240
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 742 EKDMRRMRIPVEYASHSPHIEVNDLGLLAPVPRPTGSPYISTTGDLDD-RPRDAD 801
 QY 241 YWYRNLIHQVRFSDVAQALADDGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 300
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 802 YWYRNLIHQVRFSDVAQALADDGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 861
 QY 301 DNDTRR 306
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 DB 862 SGGLR 867

RESULT 14

O52545 PRELIMINARY; PRT; 5060 AA.

AC O52545;
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Polypeptide synthase.
 GN RIFB.
 OS Amycolatopsis mediterranei (Nocardia mediterranei).
 OC Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae;
 OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae; Amycolatopsia.
 OX NCBI_TaxID=33910;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S699;
 RX MEDLINE=98174059; PubMed=9512878;
 RA August P.R., Tang L., Yoon Y.J., Ning S., Mueller R., Yu T.W.,
 Taylor M., Hoffmann D., Kim C.G., Zhang X., Hutchinson C.R.,
 RA Floss H.G.;
 RT "Biosynthesis of the anasmycin-antibiotic rifamycin: deductions from
 RT the molecular analysis of the rif biosynthetic gene cluster of
 RT Amycolatopsis mediterranei S699."
 RL Chem. Biol. 5:69-79(1998).
 DR [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S699;
 RX MEDLINE=9815773; PubMed=9497318;
 RA Kim C.G., Yu T.W., Fryhle C.B., Handa S., Floss H.G.;
 RT "3-Amino-5-hydroxybenzoic acid synthase, the terminal enzyme in the
 RT formation of the precursor of mc7N units in rifamycin and related
 RT antibiotics."
 RL J. Biol. Chem. 273:6030-6040(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S699;
 RX MEDLINE=21201076; PubMed=11278540;
 RA Yu T.-W., Mueller R., Mueller M., Zhang X., Draeger G., Kim C.-G.,
 RA Leisner E., Floss H.G.;
 RT "Mutational analysis and reconstituted expression of the biosynthetic
 RT genes involved in the formation of 3-amino-5-hydroxybenzoic acid, the
 RT starter unit of rifamycin biosynthesis in Amycolatopsis mediterranei
 RT S699.";

RL J. Biol. Chem. 276:12546-12555 (2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8699;
 RA August P.R., Tang L., Yoon Y.J., Ning S., Mueller R., Hutchinson C.R.,
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR HSP: P25715; IMLA.
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR00794; Ketoacyl-synt.
 DR InterPro: IPR03880; Pantone attach.
 DR Pfam: PF00698; Acyl transfer; 3.
 DR Pfam: PF02801; ketoacyl-synt_C; 3.
 DR Pfam: PF00550; pp-binding; 3.
 DR PROSITE; PS50075; ACP DOMAIN; 3.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 5060 AA; 526542 MW; 13A6EBE1A48A6B0 CRC64;
 Query Match 53.1%; Score 848.5; DB 2; Length 5060;
 Best Local Similarity 58.4%; Pred. No. 1.4e-51;
 Matches 180; Conservative 34; Mismatches 87; Indels 7; Gaps 5;
 QY 1 VFVPPGGGQWAGVRLAASSPVFARRMQACEALAPWYDMSVVDILRDAGDAWVERA 60
 DB 561 VWFPPGGGTQWAGVRLDSSSEVFARRIAECETALGRWVMSLTDLRGED--LIDRV 618
 QY 61 DVVQPVLFSSVWVSLAALMRSGIEPDVILGHSGEIAAAHVCGALSILDAKTVALLRSRA 120
 DB 619 DVVQPAFAVWVGLAAVWASLGVEPEAVVGHSGEIAAACVSGALSLIEDAKVVALRSQA 678
 QY 121 LAA-VKRGCGMASVPLPAQVEEQLIGERNAGRLVAAVNGPRSTAVSGDAEAVDEVLA 179
 DB 679 IAAVLGRGCGMASVPLSEEDATARTL-EPWAGRVEVAANGPSTSVIAGDAEALDEALDAL 737
 QY 180 AGTGVARRRIPVDYASHCPHVOPLRELLLELISQPSGVPFSTVEGTWL-DITLTD 238
 DB 738 DDGVRIRRRVAVDASHTRVEARDALEMLGIRAOAPEVPYSTVTGGWEDAGVLD 797
 QY 239 AAYVYRNILHOPVRFSDAVQALADGHRFVEVSPHPTLVPAIDTIEDTADVTALISLR 298
 DB 798 GGWYRNLRQVRFGRPAVALLIEGHRFVEVSAHPVLVQPINELVDTEAVVT--GTLR 855
 QY 299 RGDNDTRR 306
 DB 856 REDGGLRR 863
 RESULT 15
 ID 052789 PRELIMINARY; PRT; 5069 AA.
 AC 052789;
 DT 01-JUN-1998 (Tremblere, 06, Created)
 DT 01-JUN-1998 (Tremblere, 06, Last sequence update)
 DT 01-JUN-2002 (Tremblere, 21, Last annotation update)
 DE Rifamycin polyketide synthase, type 1.
 OS Amycolatopsis mediterranei (Nocardia mediterranei).
 OC Bacteria; Firmicutes; Actinobacteriae; Actinobacteridae;
 OC Actinomycetales; Pseudonocardiales; Pseudonocardaceae; Amycolatopsis.
 OX NCBI_TaxID=33910;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LBG A3136;
 RA Schupp T., Toupet C., Engel N., Goff S.;
 RT "Cloning and sequence analysis of the putative rifamycin polyketide
 RT synthase gene cluster from Amycolatopsis mediterranei";
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ223012; CAAL1036.1; -.
 DR HSP; P25715; IMLA.
 DR InterPro: IPR001227; Ac transferase.

DR InterPro: IPR00794; Ketoacyl-synt.
 DR InterPro: IPR03880; Pantone attach.
 DR Pfam: PF00698; Acyl transfer; 3.
 DR Pfam: PF02801; ketoacyl-synt_C; 3.
 DR Pfam: PF00550; pp-binding; 3.
 DR PROSITE; PS50075; ACP DOMAIN; 3.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 5069 AA; 527694 MW; CE545CB77C0AA4F CRC64;
 Query Match 53.1%; Score 848.5; DB 2; Length 5069;
 Best Local Similarity 58.4%; Pred. No. 1.4e-51;
 Matches 180; Conservative 34; Mismatches 87; Indels 7; Gaps 5;
 QY 1 VFVPPGGGQWAGVRLAASSPVFARRMQACEALAPWYDMSVVDILRDAGDAWVERA 60
 DB 570 VWFPPGGGTQWAGVRLDSSSEVFARRIAECETALGRWVMSLTDLRGED--LIDRV 627
 QY 61 DVVQPVLFSSVWVSLAALMRSGIEPDVILGHSGEIAAAHVCGALSILDAKTVALLRSRA 120
 DB 628 DVVQPAFAVWVGLAAVWASLGVEPEAVVGHSGEIAAACVSGALSLIEDAKVVALRSQA 687
 QY 121 LAA-VKRGCGMASVPLPAQVEEQLIGERNAGRLVAAVNGPRSTAVSGDAEAVDEVLA 179
 DB 688 IAAVLGRGCGMASVPLSEEDATARTL-EPWAGRVEVAANGPSTSVIAGDAEALDEALDAL 746
 QY 180 AGTGVARRRIPVDYASHCPHVOPLRELLLELIDISQPSGVPFSTVEGTWL-DITLTD 238
 DB 747 DDGVRIRRRVAVDASHTRVEARDALEMLGIRAOAPEVPYSTVTGGWEDAGVLD 806
 QY 239 AAYVYRNILHOPVRFSDAVQALADGHRFVEVSPHPTLVPAIDTIEDTADVTALISLR 298
 DB 807 GGWYRNLRQVRFGRPAVALLIEGHRFVEVSAHPVLVQPINELVDTEAVVT--GTLR 864
 QY 299 RGDNDTRR 306
 DB 865 REDGGLRR 872
 Search completed: June 17, 2003, 13:07:53
 Job time : 15.1889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:52:22 ; Search time 3.34793 Seconds
(without alignments)
3803.313 Million cell updates/sec

Title: US-09-914-286-3_COPY_1050_1356

Perfect score: 1598

Sequence: 1 VFVFPQGGGQMGAGVRLA.....AEDVTAIGSLRGDNDTRRP 307

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt 40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	884.5	55.4	3567	ERY2_SACER	Q03132 saccharopol
2	858	53.7	3519	OL56_STRAT	Q07017 streptomyc
3	840.5	52.6	3491	ERY1_SACER	Q03131 saccharopol
4	790	49.4	3172	ERY3_SACER	Q03133 saccharopol
5	570	35.7	2110	MCAS_MYCBO	Q02251 mycobacteri
6	493	30.9	1538	PPSB_MYCTU	Q10978 mycobacteri
7	472	29.5	1876	PPSA_MYCTU	Q10977 mycobacteri
8	448.5	28.1	1774	MSAS_PENPA	P22367 penicillium
9	322.5	20.2	1986	WA_EMENT	Q03149 emericella
10	279.5	17.5	2181	STCA_HERIN	Q12397 emericella
11	275.5	17.2	3112	FABD_HARIN	P43712 haemophilus
12	269.5	16.9	317	FABD_BACSU	P11019 bacillus su
13	256	16.0	308	FABD_SALTY	O85140 salmonella
14	254	15.9	308	FABD_ECOLI	P25715 escherichia
15	243	15.2	2505	FAS_RAT	P12785 rattus norv
16	238	14.9	2109	PKSI_ASPPA	Q12053 aspergillus
17	231.5	14.5	293	FABD_SYNY3	P13242 synecocyst
18	222.5	13.9	2511	FAS_CHICK	P12276 gallus gall
19	178.5	11.2	2504	FAS_HUMAN	P49327 homo sapien
20	125	7.8	302	FABD_MYCTU	Q10501 mycobacteri
21	106	6.6	761	CTPA_MYCTU	Q10848 mycobacteri
22	101.5	6.4	722	VATI_HALNI	Q94908 halobacteri
23	101.5	6.4	2073	FASI_SCHPO	Q94908 halobacteri
24	101	6.3	1245	NARZ_ECOLI	P19319 escherichia
25	99.5	6.2	1245	TRAB_ECOLI	P55416 thiazobium
26	97	6.1	2037	FASI_CANAL	P34731 c fatty aci
27	96	6.0	2076	FASI_YARLI	P34229 y fatty aci
28	95.5	6.0	600	GIDA_DEIRA	Q9763 delnococcus
29	95.5	6.0	693	VG50_BPMD2	O64240 mycobacteri
30	94.5	5.9	2051	FASI_YEAST	P07149 s fatty aci
31	93.5	5.9	326	YGER_ECOLI	P22052 escherichia
32	93	5.8	406	SVT2_AERPE	Q94743 aeropyrum p
33	92	5.8	533	MCPE_RHIME	O52877 rhizobium m

34	92	5.8	752	1	CTPB_MYCTU	Q10877 mycobacteri
35	90.5	5.7	629	1	SYR_PYRHO	O59147 pyrococcus
36	90.5	5.7	677	1	Y136_MYCTU	O50597 mycobacteri
37	90	5.6	421	1	TRAB_AGR75	O44351 agrobacteri
38	89.5	5.6	403	1	HMDH_HALVO	O59458 halobacteri
39	89	5.6	463	1	Y102_MYCTU	O53951 mycobacteri
40	89	5.6	2205	1	POLN_RUBVT	P13889 rubella vir
41	88	5.5	780	1	CTPA_MYCLE	P46839 mycobacteri
42	88	5.5	904	1	DPO1_MYCTU	O07700 mycobacteri
43	87.5	5.5	990	1	TNP7_ECOLI	P13694 escherichia
44	86.5	5.4	788	1	TR51_HCMVA	P09635 human cytom
45	86	5.4	418	1	Y824_DEIRA	Q9745 delnococcus

ALIGNMENTS

RESULT 1
ERY2_SACER STANDARD; PRT; 3567 AA.
ID ERY2_SACER
AC Q03132; 054096;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Erythromolide synthase, modules 3 and 4 (EC 2.3.1.94) (ORF 2) (6-deoxyerythronolide B synthase II) (DEBS 2).
GN ERYA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Pseudonocardiales; Pseudonocardaceae;
OC Saccharopolyspora.
OX NCBI_TaxID=1836;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91220065; PubMed=2024119;
RA Donadio S., Stever M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT "Modular organization of genes required for complex polyketide biosynthesis.";
RT Science 252:675-679 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RX MEDLINE=92155230; PubMed=1740151;
RA Bevil D.J., Cortes J., Haydock S.F., Leadlay P.F.;
RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme.";
RT Eur. J. Biochem. 204:39-49 (1992).
RL - CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-deoxyerythronolide B.
CC - COFACTOR: NADP. CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETINES.
CC - PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
CC - MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
CC - MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR), DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.
CC - SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
CC - SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 DR EMBL; M63677; AAA26494.1; -;
 DR EMBL; X62569; CAA44448.1; -;
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR002085; Adh zn family.
 DR InterPro; IPR000794; Ketocacyl-synt.
 DR InterPro; IPR003800; Pantne_attach.
 DR Pfam; PF00107; adh_zinc; 1.
 DR Pfam; PF00109; ketocacyl-synt; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl transf; 2.
 DR Pfam; PF02801; ketocacyl-synt C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
 DR PROSITE; PS00066; B-KETOACYL SYNTHASE; 2.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 KW Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KW Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1484 MODULE 3.
 FT DOMAIN 1485 3567 MODULE 4.
 FT DOMAIN 27 488 BETA-KETOACYL SYNTHASE 1.
 FT DOMAIN 559 884 ACYLTRANSFERASE (AT) 1.
 FT DOMAIN 1130 1301 BETA-KETOACYL REDUCTASE 1 (POSSIBLY
 NON-FUNCTIONAL).
 FT DOMAIN 1397 1467 ACYL CARRIER (ACP) 1.
 FT DOMAIN 1485 1943 BETA-KETOACYL SYNTHASE 2.
 FT DOMAIN 2013 2336 ACYLTRANSFERASE (AT) 2.
 FT DOMAIN 2383 3066 DEHYDRATASE/ENOLREDUCTASE (DH/ER).
 FT DOMAIN 3139 3322 BETA-KETOACYL REDUCTASE 2.
 FT DOMAIN 3415 3485 ACYL CARRIER (ACP) 2.
 FT ACT_SITE 202 202 THIOESTER BOND.
 FT ACT_SITE 651 651 ACYL-ENZYME INTERMEDIATE.
 FT BINDING 1430 1430 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT ACT_SITE 1661 1661 THIOESTER BOND.
 FT NP_BIND 2961 2978 ACYL-ENZYME INTERMEDIATE.
 FT NP_BIND 3142 3157 NADP (R).
 FT BINDING 3448 3448 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT CONFLICT 438 438 R -> A (IN REF. 2).
 FT CONFLICT 480 480 T -> S (IN REF. 2).
 FT CONFLICT 1241 1241 L -> F (IN REF. 2).
 FT CONFLICT 2664 2664 L -> V (IN REF. 2).
 SQ SEQUENCE 3567 AA; 374413 MW; EE6284F4738AA0C0 CRC64;

Query Match 55.4%; Score 884.5; DB 1; Length 3567;
 Best Local Similarity 57.3%; Pred. No. 4.9e-63;
 Matches 177; Conservative 41; Mismatches 88; Indels 3; Gaps 2;

QY 1 VFVPRGGGGMAGVRLASSPVFARRMQACEALAPWDMVSVDLRRDAGDAWVERA 60
 DB 560 VFLEPGGSGWAGGAGELLSSPVFAGKIRACDSMPQWKVSVDLRQAPGAGDRV 619
 QY 61 DVVQPVLFVSVVSLAALMRSGYIEPDAVLGHSGEIAAAHVCALSLDKAKTVALLSRA 120
 DB 620 DVVQPVLFVSVVSLAALMRSGYIEPDAVLGHSGEIAAAHVCALSLDKAKTVALLSRA 679
 QY 121 LAAYRGSGASVPLPAQVEQLIGERWAGLWVAANGPSTAVSGDAEAVDEVLAYCA 180
 DB 680 MRSISGSGGMAAVALGEAAVVERL-RPMQDRLSVAANGPSTAVSGEPALRAFSBDA 738
 QY 181 GTGVARIRIPVDVSHCHQPLRELLLELLGDISPSQSGVPFSTYEGMTLDTTIDAA 240
 DB 739 AEGIRVADIDVAVSHSPQIERVEELLETTGDIAPPARVTFSTVESRSMQDTLIDAR 798
 QY 241 YWYNLHQPVRFSDAVOALADGGRVFEVSPHTLPAIDTE--DTAEDVPAISGLR 298
 DB 799 YWYNLHQPVRFSDAVOALADGGRVFEVSPHTLPAIDTE--DTAEDVPAISGLR 858
 QY 299 RGDNDTRRF 307
 DB 859 RDGDLISAF 867

RESULT 2

OL56 STRAT STANDARD; PRT; 3519 AA.
 AC 007017;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Oleandomycin polyketide synthase, modules 5 and 6.
 GN ORF6.
 OS Streptomyces antibioticus.
 OC Bacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID:1890;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150470; PubMed=8107683;
 RA Swan D.G., Rodriguez A.M., Vilches C., Mendez C., Salas J.A.;
 RT "Characterization of a Streptomyces antibioticus gene encoding a type
 I polyketide synthase which has an unusual coding sequence."
 RL Mol. Gen. Genet. 242:358-362(1994).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE BIOSYNTHESIS OF THE OLEANDOMYCIN
 LACTONE RING.
 CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETHEINES.
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 DR EMBL; L09654; AAA19695.1; -;
 DR HSSP; P25715; IMLA.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketocacyl-synt.
 DR InterPro; IPR003800; Pantne_attach.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00109; ketocacyl-synt; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl transf; 2.
 DR Pfam; PF00975; Thioesterase; 1.
 DR Pfam; PF02801; ketocacyl-synt C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
 DR PROSITE; PS00066; B-KETOACYL SYNTHASE; 2.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 KW Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;
 KW Phosphopantetheine; Multifunctional enzyme; Repeat.
 FT DOMAIN 1 3519 MODULE 5.
 FT DOMAIN 32 501 MODULE 6.
 FT DOMAIN 569 890 BETA-KETOACYL SYNTHASE 1.
 FT DOMAIN 1200 1382 BETA-KETOACYL REDUCTASE 1.
 FT DOMAIN 1487 1561 ACYL CARRIER (ACP) 1.
 FT DOMAIN 1686 2186 BETA-KETOACYL SYNTHASE 2.
 FT DOMAIN 2220 2541 ACYLTRANSFERASE (AT) 2.
 FT DOMAIN 2856 3038 BETA-KETOACYL REDUCTASE 2.
 FT ACT_SITE 3141 3215 ACYL CARRIER (ACP) 2.
 FT ACT_SITE 3141 3215 ACYL-ENZYME INTERMEDIATE.
 FT BINDING 3270 3519 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT ACT_SITE 210 210 THIOESTERASE.
 FT ACT_SITE 660 660 BETA-KETOACYL SYNTHASE.
 FT NP_BIND 1203 1249 ACYL-ENZYME INTERMEDIATE.
 FT BINDING 1524 1524 NADP.
 FT ACT_SITE 1859 1859 PHOSPHOPANTHETHEINE (BY SIMILARITY).
 FT ACT_SITE 2311 2311 BETA-KETOACYL SYNTHASE.
 FT NP_BIND 2859 2905 BETA-KETOACYL SYNTHASE.
 FT BINDING 3178 3178 ACYL-ENZYME INTERMEDIATE.
 SQ SEQUENCE 3519 AA; 368561 MW; 41AE78AAAE61F86 CRC64;

Query Match 53.74; Score 858; DB 1; Length 3519;
 Best Local Similarity 57.0%; Pred. No. 6,7e-61;
 Matches 175; Conservative 34; Mismatches 94; Indels 4; Gaps 3;

QY 1 VFVFPQGGQMGWAGVRLASSPVFARMQACEALAPWDMVVDLRRDAGDAVWER 59
 DB 2221 VFVFPQGGQMGWAGVRLASSPVFARMQACEALAPWDMVVDLRRDAGDAVWER 59
 QY 60 ADVVQVPLFVSVMSLAALMRSYSGIEPDVAIGHSGEIAAAHVCGALSLKDAKTVLRSR 119
 DB 2279 VDVVQVPLFVSVMSLAALMRSYSGIEPDVAIGHSGEIAAAHVCGALSLKDAKTVLRSR 119
 QY 120 ALAAVRRGGMASVPLPAQVEQLIGRMAGRLVAAVNGPRSTAVSGDAEAVDEVLAAC 179
 DB 2339 AIAIRAGGGWAVSLPAGVRRMTL-DTYGGRVAVAAVNGPSSITVSGDVQALDELLAGC 2397
 QY 180 AGTGVARRRIPVDYASHCHVQVPLRELLGLDISQPSGVPPFSTVEGTWLDTTTLDA 239
 DB 2398 EREGVARRRIPVDYASHCHVQVPLRELLGLDISQPSGVPPFSTVEGTWLDTTTLDA 2457
 QY 240 AYWVRNHHQVFRSDAVALADGHRVFEVSPHPTLVPAIETDTEPTADVTAAIGSLRR 239
 DB 2458 GYFTNLRRTVRRQEAIVEGLVAQGMCAFVECSPPHVLPGIEQTLDAIDQAAVLSGLRR 2517
 QY 300 GDNDRTR 306
 DB 2518 DEGGDOR 2524

RESULT 3
 ERYL_SACER
 ID ERYL_SACER STANDARD; PRT; 3491 AA.
 AC 00313;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Erythronolide synthase, modules 1 and 2 (EC 2.3.1.94) (ORF 1) (6-deoxyerythronolide B synthase 1) (DEBS 1).
 GN ERYL.
 OS Saccharopolyspora erythraea (Streptomyces erythraeus).
 OC Bacteria; Actinobacteretia; Actinobacteria (class); Actinobacteridae; Actinomycetales; Pseudonocardiales; Pseudonocardiaceae;
 OC Saccharopolyspora.
 OX NCBI_TaxID=1836;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=9122005; Pubmed=2024119;
 RX Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
 RT "Modular organization of genes required for complex polyketide biosynthesis.";
 RL Science 252:675-679 (1991).
 RN [2]
 RP SEQUENCE OF 3474-3491 FROM N.A.
 RX MEDLINE=93231529; Pubmed=8386127;
 RA Donadio S., Staver M.J.;
 RT "S1136, an insertion element in the erythronycin gene cluster of Saccharopolyspora erythraea.";
 RL Gene 126:147-151 (1993).
 CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-deoxyerythronolide B.
 CC -1- COFACTOR: NADP, CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTETHEINES.
 CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
 CC -1- MISCELLANEOUS: IN EACH ORF OF ERYL TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYL SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
 CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER

CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
 CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
 CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
 CC OF THE FULL-LENGTH CHAIN.
 CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
 CC -1- SIMILARITY: CONTAINS 3 ACYL CARRIER DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M63676; AAA26493.2; -;
 CC EMBL; L07626; AAA26504.1; -;
 CC HSSP; P25715; IMLA.
 CC InterPro; IPR002198; ADH_short.
 CC InterPro; IPR001227; Ac transferase.
 CC InterPro; IPR000794; Ketoacyl-synt.
 CC InterPro; IPR003880; Pantine_attach.
 CC Pfam; PF00106; adh_short; 1.
 CC Pfam; PF00109; ketoacyl-synt; 2.
 CC Pfam; PF00550; pp-binding; 3.
 CC Pfam; PF00698; Acyl transf; 3.
 CC Pfam; PF02801; ketoacyl-synt C; 2.
 CC PROSITE; PS00012; PHOSPHOPANTETHEINE; 3.
 CC PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
 CC PROSITE; PS50075; ACP DOMAIN; 3.
 CC Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 CC KM Phosphopantetheine; Multifunctional enzyme.
 CC FT DOMAIN 1 1972
 CC FT DOMAIN 1979 3491
 CC FT DOMAIN 1 375
 CC FT DOMAIN 414 484
 CC FT DOMAIN 503 961
 CC FT DOMAIN 1030 1356
 CC FT DOMAIN 1611 1794
 CC FT DOMAIN 1888 1958
 CC FT DOMAIN 1979 2441
 CC FT DOMAIN 2507 2854
 CC FT DOMAIN 3055 3237
 CC FT DOMAIN 3334 3404
 CC FT ACT SITE 145 145
 CC FT BINDING 447 447
 CC FT ACT_SITE 677 677
 CC FT ACT_SITE 1128 1128
 CC FT NE BIND 1614 1660
 CC FT BINDING 1921 1921
 CC FT ACT_SITE 2148 2148
 CC FT ACT_SITE 2598 2598
 CC FT NE BIND 3058 3104
 CC FT BINDING 3367 3367
 CC SQ SEQUENCE 3491 AA; 365022 MW; 682BFC32C90FA8C4 CRC64;
 Query Match 52.6%; Score 840.5; DB 1; Length 3491;
 Best Local Similarity 55.1%; Pred. No. 1.7e-59;
 Matches 173; Conservative 46; Mismatches 84; Indels 11; Gaps 4;

QY 1 VFVFPQGGQMGWAGVRLASSPVFARMQACEALAPWDMVVDLRRDAGDAVWER 59
 DB 1031 VFVFPQGGQMGWAGVRLASSPVFARMQACEALAPWDMVVDLRRDAGDAVWER 1090
 QY 57 --WERADVQVPLFVSVMSLAALMRSYSGIEPDVAIGHSGEIAAAHVCGALSLKDAKTV 114
 DB 1091 LSTERVADVQVPLFVSVMSLAALMRSYSGIEPDVAIGHSGEIAAAHVCGALSLKDAKTV 1150
 QY 115 ALRSRLAAVRRGGMASVPLPAQVEQLIGRMAGRLVAAVNGPRSTAVSGDAEAVDEVLAAC 174
 DB 1151 ALRSRLAAVRRGGMASVPLPAQVEQLIGRMAGRLVAAVNGPRSTAVSGDAEAVDEVLAAC 1206
 QY 175 VLAVCAGTGVARRIPVDYASHCHVQVPLRELLGLDISQPSGVPPFSTVEGTWLD 233

```

Db 1207 LVASCTEICTRACKLAVDASHSHVETIRDALHAELGEDHPHPCGVFPFSTVYGRWTQ 1266
Cc 1266
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.lsb-sib.ch/announce/
Cc or send an email to license@lsb-sib.ch).
Cc -----
Cc EMBL; X56107; CAA39583.1; -
Cc EMBL; M63677; AAA26495.1; -
Cc EMBL; X62569; CAA44449.1; -
Cc HSP; P00101; 1CCH
Cc InterPro; IPR002196; ADH_short.
Cc InterPro; IPR001227; Ac transferase.
Cc InterPro; IPR000794; ketoacyl-synt.
Cc InterPro; IPR003880; Pantine attach.
Cc InterPro; IPR001031; Thioesterase.
Cc Pfam; PF00106; adh_short; 1.
Cc Pfam; PF00109; ketoacyl-synt; 2.
Cc Pfam; PF00550; pp-binding; 2.
Cc Pfam; PF00698; Acyl transf; 2.
Cc Pfam; PF00975; Thioesterase; 1.
Cc Pfam; PF02801; ketoacyl-synt C; 2.
Cc PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
Cc PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
Cc PROSITE; PS50075; ACP DOMAIN; 2.
Cc Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
Cc Phosphopantetheine; Multifunctional enzyme.
Cc -----
Cc MODULE 5.
Cc MODULE 6.
Cc BETA-KETOACYL SYNTHASE 1.
Cc ACYLTRANSFERASE (AT) 1.
Cc BETA-KETOACYL REDUCTASE 1.
Cc ACYL CARRIER (ACP) 1.
Cc BETA-KETOACYL SYNTHASE 2.
Cc ACYLTRANSFERASE (AT) 2.
Cc BETA-KETOACYL REDUCTASE 2.
Cc ACYL CARRIER (ACP) 2.
Cc THIOESTERASE.
Cc ACYL-ENZYME INTERMEDIATE.
Cc NADP.
Cc PHOSPHOPANTETHEINE (BY SIMILARITY).
Cc THIOESTER BOND.
Cc ACYL-ENZYME INTERMEDIATE.
Cc NADP.
Cc PHOSPHOPANTETHEINE (BY SIMILARITY).
Cc MISSING (IN REF. 2).
Cc A -> R (IN REF. 2).
Cc P -> R (IN REF. 2).
Cc PEPRNSLRDGFILATRASMEHRA -> ASRGTCAIPVS
Cc RMPAPAPMO (IN REF. 1).
Cc A -> R (IN REF. 2).
Cc W -> W (IN REF. 2).
Cc E -> D (IN REF. 2).
Cc R -> G (IN REF. 2).
Cc GPNSP -> ARTR (IN REF. 2).
Cc R -> G (IN REF. 2).
Cc R -> A (IN REF. 2).
Cc MISSING (IN REF. 2).
Cc AHR -> GIT (IN REF. 2).
Cc R -> ROR (IN REF. 2).
Cc R -> RELPYEPOROR (IN REF. 1).
Cc GVAAYH -> VLLSARD (IN REF. 2).
Cc RTMPLERLA -> ARTRMSR (IN REF. 2).
Cc MISSING (IN REF. 1).
Cc L -> V (IN REF. 2).
Cc A -> R (IN REF. 2).
Cc MISSING (IN REF. 2).
Cc AA -> RR (IN REF. 2).
Cc LCGRE -> STER (IN REF. 2).
Cc MISSING (IN REF. 2).
Cc G -> R (IN REF. 2).
Cc -----
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.lsb-sib.ch/announce/
Cc or send an email to license@lsb-sib.ch).
Cc -----
Cc EMBL; X56107; CAA39583.1; -
Cc EMBL; M63677; AAA26495.1; -
Cc EMBL; X62569; CAA44449.1; -
Cc HSP; P00101; 1CCH
Cc InterPro; IPR002196; ADH_short.
Cc InterPro; IPR001227; Ac transferase.
Cc InterPro; IPR000794; ketoacyl-synt.
Cc InterPro; IPR003880; Pantine attach.
Cc InterPro; IPR001031; Thioesterase.
Cc Pfam; PF00106; adh_short; 1.
Cc Pfam; PF00109; ketoacyl-synt; 2.
Cc Pfam; PF00550; pp-binding; 2.
Cc Pfam; PF00698; Acyl transf; 2.
Cc Pfam; PF00975; Thioesterase; 1.
Cc Pfam; PF02801; ketoacyl-synt C; 2.
Cc PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
Cc PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
Cc PROSITE; PS50075; ACP DOMAIN; 2.
Cc Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
Cc Phosphopantetheine; Multifunctional enzyme.
Cc -----
Cc MODULE 5.
Cc MODULE 6.
Cc BETA-KETOACYL SYNTHASE 1.
Cc ACYLTRANSFERASE (AT) 1.
Cc BETA-KETOACYL REDUCTASE 1.
Cc ACYL CARRIER (ACP) 1.
Cc BETA-KETOACYL SYNTHASE 2.
Cc ACYLTRANSFERASE (AT) 2.
Cc BETA-KETOACYL REDUCTASE 2.
Cc ACYL CARRIER (ACP) 2.
Cc THIOESTERASE.
Cc ACYL-ENZYME INTERMEDIATE.
Cc NADP.
Cc PHOSPHOPANTETHEINE (BY SIMILARITY).
Cc THIOESTER BOND.
Cc ACYL-ENZYME INTERMEDIATE.
Cc NADP.
Cc PHOSPHOPANTETHEINE (BY SIMILARITY).
Cc MISSING (IN REF. 2).
Cc A -> R (IN REF. 2).
Cc P -> R (IN REF. 2).
Cc PEPRNSLRDGFILATRASMEHRA -> ASRGTCAIPVS
Cc RMPAPAPMO (IN REF. 1).
Cc A -> R (IN REF. 2).
Cc W -> W (IN REF. 2).
Cc E -> D (IN REF. 2).
Cc R -> G (IN REF. 2).
Cc GPNSP -> ARTR (IN REF. 2).
Cc R -> G (IN REF. 2).
Cc R -> A (IN REF. 2).
Cc MISSING (IN REF. 2).
Cc AHR -> GIT (IN REF. 2).
Cc R -> ROR (IN REF. 2).
Cc R -> RELPYEPOROR (IN REF. 1).
Cc GVAAYH -> VLLSARD (IN REF. 2).
Cc RTMPLERLA -> ARTRMSR (IN REF. 2).
Cc MISSING (IN REF. 1).
Cc L -> V (IN REF. 2).
Cc A -> R (IN REF. 2).
Cc MISSING (IN REF. 2).
Cc AA -> RR (IN REF. 2).
Cc LCGRE -> STER (IN REF. 2).
Cc MISSING (IN REF. 2).
Cc G -> R (IN REF. 2).
Cc -----
Cc This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc -----

```

Query Match	Similarity	Score	DB 1	Length
Beat Local	56.0%	Pred. No. 1.8e-55;		
Matches 169;	Conservative 39;	Mismatches 84;	Indels 10;	Gaps 5
FT	CONFLICT	1760	1601	V -> L (IN REF. 2).
FT	CONFLICT	1724	1725	LP -> FA (IN REF. 2).
FT	CONFLICT	1732	1732	Q -> L (IN REF. 2).
FT	CONFLICT	1739	1743	GAEAG -> ARRA (IN REF. 2).
FT	CONFLICT	1762	1762	T -> S (IN REF. 2).
FT	CONFLICT	2252	2252	D -> DGAD (IN REF. 2).
FT	CONFLICT	2275	2277	GGP -> AVA (IN REF. 2).
FT	CONFLICT	2420	2408	G -> GR (IN REF. 2).
FT	CONFLICT	2420	2421	LA -> S (IN REF. 2).
FT	CONFLICT	2443	2444	NA -> TH (IN REF. 2).
FT	CONFLICT	2596	2596	A -> G (IN REF. 2).
FT	CONFLICT	2609	2609	P -> A (IN REF. 2).
FT	CONFLICT	2715	2722	RAEERA -> AVKAVR (IN REF. 1).
FT	CONFLICT	2754		D -> E (IN REF. 2).
SO	SEQUENCE	3172 AA;	331474 MW;	DBD5094E77DD5F CRC64;
Query Match	49.4%	Score 790;	DB 1;	Length 3172;
Beat Local	56.0%	Pred. No. 1.8e-55;		
Matches 169;	Conservative 39;	Mismatches 84;	Indels 10;	Gaps 5
QY	3	VFPFGGGWAGNGVRLASSPVFARRMOACEBALAPWYDMSVVDILRRDAGDAVERADV	62	
DB	557	VFPFGGAGWQMGARDDLRESQVFADISIRDCERBALPHYDMSVLTDL--SGARPDRVDV	613	
QY	63	VQPVLFSTWVSLAALMRSYGYTPDVLVGHSGEILAAHVCGSLSKDAKAYALSRALA	122	
DB	614	VQPAFAFAWVSLAALMRSHGVPAVVGHSQGEILAAHVAGALTLEDAKVAVRSVLR	673	
QY	123	AVRGGGASVYLPFAOEVEOLIGERWAGRLWVAANGPRS-TAVSGDAVDEVLYAYCG	181	
DB	674	RLGGGGGASRLGLTQQAERIG-RFAGALSLASNGPRSVVYVAGSGPLDELTAECBA	732	
QY	182	TGVARRIPVDYASHCPVQPLREELBELLDGISPOSGVPFFSTVEGTWLDITTLDAAY	241	
DB	723	EAHKARRIPVDYASHSPQVESIRELLELDELIGISVSDVNLVSTTGQPIDTATMDRAY	792	
QY	242	WYRNLMHGVRRSDVAQALADDGHRVFEVSPHPTVPAIEDTTEADVTA---ISL	297	
DB	793	WYANLRGVRRFQDARQLAEGFDFVEVSPHPTVGVIE-ATLDSALPADAGACVGTLL	851	
QY	298	RR 299		
DB	852	RR 853		
RESULT 5				
MCAS_MYCBO	STANDARD;	PRT;	2110 AA.	
ID	MCAS_MYCBO			
AC	002251;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	Mycocerosic acid synthase.			
GN	MAS.			
OS	Mycobacterium bovis.			
OC	Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;			
OC	Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
CC	NCBI_TaxID=1765;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE=92406887; PubMed=1527058;			
RA	Mathur M., Kolatukudy P.E.;			
RT	"Molecular cloning and sequencing of the gene for mycocerosic acid			
RT	synthase, a novel fatty acid elongating multifunctional enzyme, from			
RT	Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guérin.";			
RL	J. Biol. Chem. 267:19388-19395(1992).			
CC	-1- FUNCTION: CATALYZES THE ELONGATION OF N-FATTY ACYL-COA WITH			
CC	METHYLMALONYL-COA (NOT MALONYL-COA) AS THE ELONGATING AGENT TO			
CC	FORM MYCOCEROSYL LIPIDS.			
CC	-1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETINE			
CC	-1- SUBUNIT: HOMODIMER WHOSE MONOMERS PROBABLY HAVE A HEAD TO TAIL			
CC	ARRANGEMENT.			
CC	-1- SUBCELLULAR LOCATION: Membrane-associated.			

CC	-1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC	or send an email to license@ebi.ac.uk).
CC	-----
DR	EMBL; M95808; AAA25369.1; .
DR	PIR; B44110; B44110.
DR	HSSP; P73283; IBSM.
DR	InterPro; IPRO01227; Ac transferase.
DR	InterPro; IPRO02085; Adh_zn_family.
DR	InterPro; IPRO00794; ketoacyl-synt.
DR	InterPro; IPRO03880; Pantane_attach.
DR	Pfam; PF00107; adh_zinc; 1.
DR	Pfam; PF00109; ketoacyl-synt; 1.
DR	Pfam; PF00550; pp-binding; 1.
DR	Pfam; PF00698; Acyl_transfer; 1.
DR	Pfam; PF02801; ketoacyl-synt_C; 1.
DR	PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR	PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR	PROSITE; PSS0075; ACP_DOMAIN_1.
KW	Fatty acid biosynthesis; Multifunctional enzyme; Phosphopantetheine;
KW	transferase; Hydrolyases; Oxidoreductase; Ligase; NADP; Membrane.
FT	DOMAIN 1 430
FT	DOMAIN 533 852
FT	DOMAIN ? ?
FT	DOMAIN ? ?
FT	ACT_SITE 177 177
FT	ACT_SITE 623 623
FT	NP_BIND 1561 1578
FT	NP_BIND 1765 1780
FT	BINDING 2059 2059
SQ	SEQUENCE 2110 AA; 225577 MW; CC658215D7155300 CRC64;
Query March	35.7%; Score 570; DB 1; Length 2110;
Best Local Similarity	41.3%; Pred. No. 6,4e-38;
Matches 124; Conservative 49; Mismatches 123; Indels 4; Gaps 2;	
Qy	1 VWFPGGGGGMAGVLLASSPVFARRMOACEBALAPWDWSVDILRRDAGDAWERA 60
Dd	534 VWVFSGGGGWMAWGTLASBPFAATIAKLEPVIAAESGFSTEAR---TAQQVTGI 590
Qy	61 DVVQPVLFSPVNSLAALMSYGTIEPAVLGHSGEITAAAHVGALSLKDAKTVALRSR 119
Dd	591 DKVQPAFAAQAVALAAMTEQTGYGPBAVVGHSHGESAAAVVAGALSLEDPAARYICRRSK 650
Qy	120 ALAAVRGRGGMASPLPAQVEEDLIGRWAGRMLVAAVNPPRAVSAGDAEVADEVAYC 179
Dd	651 LMRIRIAGAMGSVELDPKAKVNSELARGIDDDVAVSVASPOSTVIGGSTDYRDLIARW 710
Qy	180 AGTVGARRRIPVDYASHCHPAOVLREBELLGLDISPQSGVPFFSTEGSTWLDTTLDA 239
Dd	711 EQRDVMAREVAIVDVASHSPQVDPILDDLAAALADIAMPVKVYYSAITLPDRPQPDCG 770
Qy	240 AYWRNLHQPVRFSDAVOALADGHRVFEVSPHPTLVPAIEDTTEDTAEDVTAIGSLR 299
Dd	771 AYWMNDLRNTVDFEAALAAQAMEDGYRVFADVSPHLTLHAVEQTGRSLDWSVAALAGMR 830
RESULT 6	
ID	PSB MYCTU STANDARD; PRT; 1538 AA.
AC	Q10978; O53234;
DT	01-OCT-1996 (Rel. 34, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DE	16-OCT-2001 (Rel. 40, Last annotation update)
DN	Phenolphthalein synthase polyketide synthase ppsb.
EN	PPSB OR RV2392 OR MT3002 OR MTCYC38.21 OR MY0011.01.

OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxId=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornby T., Jagels K., Kirogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Suleron J.E., Taylor K., Whitehead S., Barrell B.G.,
 RA "deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLPTHIICEROL
 CC SYNTHESIS.
 CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTETHEINE (By
 CC similarity).
 CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL021070; CA15929.1;
 CC EMBL: AE007122; AK47329.1;
 CC TIGR: M33002;
 DR TuberculList; RV2932;
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Ppantne_attach.
 DR Pfam; PF00109; ketoacyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR Pfam; PF00658; Acyl-transfer; 1.
 DR Pfam; PF02801; ketoacyl-synt C; 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; FALSE_NEG.
 DR PROSITE; PS50075; ACP_DOMAIN; 1.
 DR Hypothetical protein; Multifunctional enzyme; Oxidoreductase;
 DR Transferrase; NADP; Phosphopantetheine; Complete proteome.
 KM TRANSFERASE; NADP; Phosphopantetheine; Complete proteome.
 FT BINDING 1458 1458 ACYL CARRIER (ACP).
 FT BINDING 1458 1458 PHOSPHOPANTETHEINE (POTENTIAL).
 SQ SEQUENCE 1538 AA; 162527 MW; B55E2A2042AD00CC CRC64;
 Query Match 30.9%; Score 493; DB:1; Length 1538;
 Best Local Similarity 35.7%; Pred. No. 7.2e-32;
 Matches 114; Conservative 56; Mismatches 129; Indels 20; Gaps 6;
 QY 1 VFVFGGGGQWAGVRLASSPVFARMQACEALAPWDMVVDILRRDADGAWERA 60
 DB 561 VFVYSGSGQWAGVRLADDEPAFAVALEPVFEQAGFSLRDVT--ATGKEIVGI 617
 QY 61 DVVQPVLFVWVSLAALMRSYGIPEPAVLGHSQGEIAAAHVCGALSLKDAKTVALSRA 120

DB 618 EQQLGLIGLQGLUTELTMSYGVQPDVLVGHSGVEVAANVAGALPAEGLRVATARRAL 677
 QY 121 LAAVRGGGASVPLPAQVEQLIGRMAGRLVVAANVGRSTAVSGDAEVADEVAYCA 180
 DB 678 MAPLSGGGVALGLDPAALTEALID--YQVTVGIVNSPRQVIVGPEQIDELIARV 735
 QY 181 GTGVARRRIPVDYASHCPHVPPLRELELLELGDISPQSGVPPSTVEGTWLDPTT---L 237
 DB 736 AQNRPRSRVAVIEVAPNPNMADLQPMRSELADLTRPTIGIST---TYADLHQPIF 792
 QY 238 DAAVYRNILHQVRFSDAVQAL---ADGHRVFEVSPHPTLVPAIEDTTD-----TA 288
 DB 793 DAEHWATNMKNPVRFOGALASAGSGADGAVHFTFISAPHLTQALADLTDEAHRPTKSA 852
 QY 289 EDVTAIGSLRGNDTRRF 307
 DB 853 AKYLSIGTLQRDADDTVTF 871
 RESULT 7
 PPISA_MYCTU STANDARD; PRT; 1876 AA.
 ID PPISA_MYCTU
 AC Q10977;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenolphthiocerol synthetase polyketide synthase ppisa.
 GN PPISA OR RV2931 OR M33000 OR MTC338.20.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxId=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornby T., Jagels K., Kirogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Suleron J.E., Taylor K., Whitehead S., Barrell B.G.,
 RA "deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLPTHIICEROL
 CC SYNTHESIS.
 CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES (By
 CC similarity).
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; 274697; CAA98988.1; -
 DR EMBL; AE007122; AAK47328.1; -
 DR TIGR; MT3000; -
 DR Tuberculester; RV2931; -
 DR InterPro; IPR001227; AC_transferase.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Pantane_attach.
 DR Pfam; PF00109; ketoacyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl_transfer; 1.
 DR Pfam; PF02801; ketoacyl-synt C; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; 1.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 KM Multifunctional enzyme; Oxidoreductase; Transferase; NADP; Repeat;
 KM Phosphopantetheine; Complete proteome.
 FT DOMAIN 7 80 ACYL CARRIER (ACP) 1.
 FT BINDING 1764 1833 ACYL CARRIER (ACP) 2.
 FT BINDING 1498 1503 NADP (POTENTIAL).
 FT BINDING 62 62 PHOSPHOPANTETHEINE (POTENTIAL).
 FT ACT_SITE 273 273 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT_SITE 720 720 MALONYLTRANSFERASE (BY SIMILARITY).
 FT BINDING 1796 1796 PHOSPHOPANTETHEINE (POTENTIAL).
 FT CONFLICT 624 624 D -> E (IN REF. 2).
 FT CONFLICT 877 877 R -> H (IN REF. 2).
 FT CONFLICT 1323 1323 G -> S (IN REF. 2).
 SQ SEQUENCE 1876 AA; 198834 MW; D9783DBD48792110 CRC64;

Query Match 29.5%; Score 472; DB 1; Length 1876;
 Best Local Similarity 34.9%; Pred. No. 4.5e-30;
 Matches 110; Conservative 54; Mismatches 135; Indels 16; Gaps 6;

QY 1 VFVPGGQGMAGVLLASSPFAARMQACEBALAPWDMVSVDLRRDAGDAWERA 60
 DB 632 VFVPGGQGMAGVLLASSPFAARMQACEBALAPWDMVSVDLRRDAGDAWERA 60
 QY 61 DVVQPVLFVWVSLAALMRSYGIIEPDVAVLGHSGGEIAAAHVCALSLKDAKTVALSRA 120
 DB 689 EQQLGLIGMQLATLMCGYVRPDLVIGHSGEVAAVAAGLTPAEGARATATSRLL 748
 QY 121 LAAVRGGMASVPLPAQEVQOLIGERWAGRLVVAANGPRSTAVSGDAEAVDEVALYCA 180
 DB 749 MAPLSCGGMALLLELAPTEALAD--FPQVTLGINSRQVTIAGPTQIDELIARVR 806
 QY 181 GTGVARRIPVDYASHCPHQPUREBELLELGDISPPQSGPPFSTVEGTMLDTT--L 237
 DB 807 AQNRFAARVIEVAPHPNPAMDALQPARSELADLTPTPTTIGIIST--TYADLHTQPVF 863
 QY 238 DAAVWYRNLIQVRFSDAVOAL--ADDGHRVFEVSPHPTLVPAIEDTTEADV--T 292
 DB 864 DAHEWAMNMPVRFQQAIALSAGSAGACGAVHTFIEISAHPLTQALIDTLHSAQPGARYT 923
 QY 293 AIGSLARGNDTRRF 307
 DB 924 SLGTGLQRTDDVVTF 938
 RESULT 8
 MSAS_PENPA STANDARD; PRT; 1774 AA.
 AC P2357;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 6-methylsalicylic acid synthase (EC 2.3.1.-) (MSAS).
 OS Penicillium patulum (Penicillium griseofulvum).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiomycetes; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
 NX NCBI_TaxID=5078;
 [1]
 RF SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=DSM 62862;
 RX MEDLINE=91006137; PubMed=2209605;

RA Beck J., Ripka S., Siegner A., Schiltz E., Schweizer E.;
 RT "The multifunctional 6-methylsalicylic acid synthase gene of
 RT Penicillium patulum. Its gene structure relative to that of other
 RT polycyclic synthases";
 RT Eur. J. Biochem. 192:487-498 (1990).
 CC -1- FUNCTION: This multifunctional enzyme is a polycyclic synthase.
 CC It catalyzes a total of 11 steps by seven different component
 CC enzymes, in the biosynthesis of the antibiotic patulin.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + 3 malonyl-CoA + NADPH =
 CC 6-methylsalicylic acid + NADP(+) + 3 CoA + 3 CO(2) + H(2)O.
 CC -1- PATHWAY: Patulin biosynthesis.
 CC -1- SUBUNIT: HOMOMULTIMER.
 CC -1- INDUCTION: IN THE LATE LOGARITHMIC GROWTH PHASE.
 CC -1- SIMILARITY: WITH MOST POLYKETIDE SYNTHASES, SPECIALLY WITH RAT
 CC FATTY ACID SYNTHASE, AND WITH OTHER ENZYMES SUCH AS LIPASES AND
 CC THIOLEASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC DR EMBL; X55776; CAA39295.1; -
 CC PIR; S13178; S13178.
 DR InterPro; IPR001227; AC_transferase.
 DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Pantane_attach.
 DR Pfam; PF00109; ketoacyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR Pfam; PF00698; Acyl_transfer; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; 1.
 DR PROSITE; PS50075; ACP DOMAIN; 1.
 KM Multifunctional enzyme; Oxidoreductase; Antibiotic biosynthesis;
 KM Transferase; NADP; Phosphopantetheine.
 FT DOMAIN 186 238 ACYLTRANSFERASE (AT).
 FT DOMAIN 642 676 ACETYL/MALONYL TRANSFERASES.
 FT DOMAIN 1403 1450 2-OXOACYL REDUCTASE.
 FT BINDING 1700 1769 ACYL CARRIER (ACP).
 FT BINDING 1419 1424 NADP (POTENTIAL).
 FT ACT_SITE 204 204 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT_SITE 653 653 MALONYLTRANSFERASE (BY SIMILARITY).
 FT BINDING 1732 1732 PHOSPHOPANTETHEINE (BY SIMILARITY).
 SQ SEQUENCE 1774 AA; 190732 MW; 05ED5DD10863F938 CRC64;

Query Match 28.1%; Score 448.5; DB 1; Length 1774;
 Best Local Similarity 38.2%; Pred. No. 3.3e-28;
 Matches 116; Conservative 46; Mismatches 127; Indels 15; Gaps 7;

QY 1 VFVPGGQGMAGVLLASSPFAARMQACEBALAPWDMVSVDLRRDAGDAWERA 60
 DB 567 VFVPGGQGMAGVLLASSPFAARMQACEBALAPWDMVSVDLRRDAGDAWERA 60
 QY 61 DVVQPVLFVWVSLAALMRSYGIIEPDVAVLGHSGGEIAAAHVCALSLKDAKTVALSRA 120
 DB 622 DRVQILTYWQIGLSALDSNGITPQAVIGHSGEIAASVAGALSAPBAGALVTTRALL 661
 QY 121 LAAVRGGMASVPLPAQEVQOLIGERWAGRLVVAANGPRSTAVSGDAEAVDEVALYCA 180
 DB 682 YRVYMGKGMILVNLDSAEFEELTGR--SDLVVAIDSSRSGCVAGDKEVAFETBAUK 739
 QY 181 GTGVARRIPVDYASHCPH---VQPLREBELLELGDISPPQSGPPFSTVEGTMLDTT 236
 DB 740 ARGVKTFTVSDIAFHSPTLGLVDPDLRDLVLETLTPVSP---NVKLSTALADPPGQDL 796
 QY 237 LDAVWYRNLIQVRFSDAVOALADDGHRVFEVSPHPTLVPAIEDTTEADV--EDVTAIG 295
 DB 797 RDVEYVAGNKNVNRRLTSAKAAVDEGYRFLFLEVSTHPVSHSINELTMDAGMEDRAVIP 856

Qy 296 SLRR 239
Db 857 TLRR 860

RESULT 9

WA_EMENT STANDARD; PRT; 1986 AA.
ID WA_EMENT
AC 003149;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Confidial green pigment synthase (EC 2.3.1.-).
GN WA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emericella.
OX NCBI_TaxId=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9101122; PubMed=1465094;
RA Mayorga M.E., Timberlake W.E.,
RT "The developmentally regulated Aspergillus nidulans wa gene encodes a
RT polypeptide homologous to polyketide and fatty acid synthases.";
RL Mol. Genet. 235:205-212(1992).
CC -1- FUNCTION: THIS PROTEIN CONDENSES CARBON UNITS TO FORM AN
CC INTERMEDIATE YELLOW POLYKETIDE PIGMENT THAT IS POLYMERIZED
CC BY CONFIDIAL LACCASE TO FORM THE GREEN PIGMENT IN MATURE
CC ASEXUAL SPORES (CONIDIA).
CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES
CC (POTENTIAL).
CC -1- PATHWAY: Confidial green pigment biosynthesis.
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X65866; CAA6695.1; -
DR PIR; S28353; S28353.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketocacyl-synt.
DR InterPro; IPR003880; ppanne_attach.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00109; ketocacyl-synt; 1.
DR Pfam; PF00550; pp-binding; 2.
DR Pfam; PF00698; Acyl transf; 1.
DR Pfam; PF00975; Thioesterase; 1.
DR Pfam; PF02801; ketocacyl-synt C; 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR PROSITE; PS50075; ACP DOMAIN; 2.
KM Transferase: Phosphopantetheine: Multifunctional enzyme: Repeat.
FT DOMAIN 529 582
FT DOMAIN 991 1024
FT DOMAIN 1650 1719
FT DOMAIN 1772 1841
FT ACT_SITE 548 548
FT ACT_SITE 1001 1001
FT BINDING 1682 1682
FT BINDING 1804 1804
FT SEQUENCE 1986 AA; 216634 MW; 74EF0940FF40EB9A CRC64;
Query Match 20.2%; Score 322.5; DB 1; Length 1986;
Best Local Similarity 29.6%; Pred. No. 5; 6e-18;
Matches 95; Conservative 54; Mismatches 137; Indels 35; Gaps 9;

RESULT 10

STCA_EMENT STANDARD; PRT; 2181 AA.
ID STCA_EMENT
AC Q12397;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative sterigmatocystin biosynthesis polyketide synthase (PKS).
GN STCA OR PKST.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emericella.
OX NCBI_TaxId=5072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 26;
RX MEDLINE=96202293; PubMed=8643646;
RA Brown D.W., Yu J.-H., Kelkar H.S., Fernandes M., Nesbitt T.C.,
RA Keller N.P., Adams T.H., Leonard T.J.;
RT "Twenty-five coregulated transcripts define a sterigmatocystin gene
RT cluster in Aspergillus nidulans.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:1418-1422(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=FGSC 4;
RX MEDLINE=95370159; PubMed=7642507;
RA Yu J.-H., Leonard T.J.;
RT "Sterigmatocystin biosynthesis in Aspergillus nidulans requires a
RT novel type I polyketide synthase.";
RL J. Bacteriol. 177:4792-4800(1995).
CC -1- FUNCTION: INVOLVED IN THE SYNTHESIS OF THE POLYKETIDE NUCLEUS OF
CC STERIGMATOCYSTIN FROM HEXANOYL-CoA AND SEVEN MALONATES.
CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES
CC (POTENTIAL).
CC -1- PATHWAY: Sterigmatocystin biosynthesis; first step.
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U34740; AAC49191.1; -
DR EMBL; L39121; AAA61586.1; -
DR HSSP; P25715; TMLA.

DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR000794; ketoacyl-synt.
 DR InterPro: IPR003880; Ppancne attach.
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00109; ketoacyl-synt; 1.
 DR Pfam: PF00550; pp-binding; 2.
 DR Pfam: PF00698; Acyl transfer; 1.
 DR Pfam: PF00975; Thioesterase; 1.
 DR Pfam: PF02801; ketoacyl-synt C; 1.
 DR POSITE: PSS0075; ACP DOMAIN 2.
 KM Transferase; Acyltransferase; phosphopantetheine; Repeat;
 KM Multifunctional enzyme.
 FT DOMAIN 383 814 BETA-KETOACYL SYNTHASE.
 FT DOMAIN 884 1209 ACYL/MALONYL TRANSFERASES.
 FT DOMAIN 1706 1777 ACYL CARRIER (ACP) 1.
 FT DOMAIN 1830 1901 ACYL CARRIER (ACP) 2.
 FT DOMAIN ? 2181 THIOESTERASE.
 FT ACT SITE 552 552 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT SITE 978 978 ACYL/MALONYL TRANSFERASES (BY SIMILARITY).
 FT BINDING 1738 1738 PHOSPHOPANTHEINE (BY SIMILARITY).
 FT BINDING 1862 1862 PHOSPHOPANTHEINE (BY SIMILARITY).
 FT ACT SITE 2028 2028 THIOESTERASE (BY SIMILARITY).
 SQ SEQUENCE 2181 AA; 238831 MW; 5A3E5712AA9AD942 CRC64;

Query Match 17.5%; Score 279.5; DB 1; Length 2181;
 Best Local Similarity 26.9%; Pred. No. 1.9e-14;
 Matches 82; Conservative 62; Mismatches 148; Indels 13; Gaps 7;

QY 1 VFPPGCGGAGMGVRLASSPVFARRMOCCEBALAPWVDVDIRRDAGDAVWER 59
 DB 885 VIACGSGSQYTGMLVNSYPTFRSDLERPDOLARSYFPFLLEVYTSKPGVDSMEDL 944
 QY 60 ADV-VOPVLFVWVSLAALRSYGIPEADVIGHSGEIIAAHVCALSLDKAAKTVALRS 118
 DB 945 LPVIVQALVSLAALRSYGIPEADVIGHSGEIIAAHVCALSLDKAAKTVALRS 1004
 QY 119 RALAAGRGGMASVPLPAQEV--QLIGERWAGRLVAAVNGRSTAVSGDAEAVDEV 176
 DB 1005 KLGQRCQRGTHMLAVRASPVTLCEVLA--SNCVACHNGNDVLSGPKAEWNLQ 1061
 QY 177 AYACGTGVRARRIPVDYASHCPHVQPLREBELLELDISPOSGVPPFSTVEGTWLDTT- 235
 DB 1062 NSLSATGCIKGTLLKLPAPFASQAQVQPLLEFKNVARSVTFHKKQIPVLSPLVVIDEKG 1121
 QY 236 TLDAAYWYRLHQPVRSDAVQALAD---GHRVPEVSPHPTLVPAIDTE-DPAED 230
 DB 1122 TVDPVYLARHCRBPVWVSLAALRSYGIPEADVIGHSGEIIAAHVCALSLDKAAKTVALRS 1181
 QY 291 VTAIG 295
 DB 1182 LPTLG 1186

RESULT 11
 FADB HAEIN STANDARD; PRT; 312 AA.
 ID FADB HAEIN STANDARD; PRT; 312 AA.
 AC P43712;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39) (MCT).
 OS FADB OR H10156.
 OC Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KM20 / ATCC 51907;
 RX MEDLINE=55350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodex A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Pine L.D., Fritchman J.L., Funtman J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Frazer C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC Malonyl-[acyl-carrier protein].
 CC -1- PATHWAY: Fatty acid biosynthesis.
 CC -1- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH
 CC TRANSACYLASE ACTIVITY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U32701; AAC21825.1; -.
 DR HSSP: P25715; IMLA.
 DR TIGR: H10156; -.
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR004410; FADB.
 DR Pfam: PF00698; Acyl transfer; 1.
 DR TIGRPFAM: TIGR00128; fadb; 1.
 KM Fatty acid biosynthesis; transferase; Complete proteome.
 FT ACT SITE 93 93 BY SIMILARITY.
 FT ACT SITE 202 202 BY SIMILARITY.
 SQ SEQUENCE 312 AA; 33424 MW; 6261622B834AB878 CRC64;

Query Match 17.2%; Score 275.5; DB 1; Length 312;
 Best Local Similarity 28.5%; Pred. No. 3.9e-15;
 Matches 91; Conservative 47; Mismatches 146; Indels 35; Gaps 12;

QY 3 VFPPGCGGAGMGVRLASSPVFARRMOCCEBALAPWVDVDIRRDAGDAVWER 62
 DB 7 VFPPGCGSQYTGMLVNSYPTFRSDLERPDOLARSYFPFLLEVYTSKPGVDSMEDL 61
 QY 63 VOPVLFVWVSLAALRSY--GIEPDVIGHSGEIIAAHVCALSLDKAAKTVALRSRA 120
 DB 62 TOPALAAVAITRWKKEKRPQIKPEVMHSGEISALVCAGVLPQDHIKVELRGKL 121
 QY 121 L-AAVGRGGM-ASVPLPAQEVQLIGERWAGRLVAAV--NGPRSTAVSGDAEAVDEV 175
 DB 122 MQDAVPEGGMAYAIIGLNEALINACKQABEGV--VSAVNFNSPGQVVIAGAKAVERA 180
 QY 176 LAYCAGTV-RARRIVDYASHCPHVQPLREBELLELDISPOSGVPPFSTV-----E 228
 DB 181 AALCKEAGAKRALPLVAVSPSHCALMKPABQALVLTENIQINTPTISVANNVDAEET 240
 QY 229 GTWLDITTLDAAYWYRLHQPVRSDAVQALADDDGHRVPEVSPHPTLVPAIDTEEDTA 288
 DB 241 GTEIRTAIV-----RLTSPVMTETVERKMDGVLAEVGPGRVLT-----NGLTKRIY 290
 QY 289 EDVTAIGSLRGNDTRRF 307
 DB 291 GDLQATISV-----NDVASF 304

RESULT 12
 FADB BACSU STANDARD; PRT; 317 AA.
 ID FADB BACSU STANDARD; PRT; 317 AA.
 AC P71019; O34463;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39) (MCT).

GN FADB.
 OC Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=96326321; PubMed=9759840;
 RA Morbidon H.R., de Mendoza D., Cronan J.E. Jr.;
 RT "Bacillus subtilis acyl carrier protein is encoded in a cluster of
 lipid biosynthesis genes";
 RL J. Bacteriol. 178:4794-4800 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98195738; PubMed=9534248;
 RA Foulger D., Errington J.;
 RT "A 28 kbp segment from the spovm region of the Bacillus subtilis 168
 genome";
 RL Microbiology. 144:801-805 (1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 Azevedo V., Berrero M.G., Bessieres P., Bolotin A., Borchert S.,
 Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 Brouillet S., Brunsch C.V., Caldwell B., Capuano N.J., Carter N.M.,
 Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 Danizov F., Devine K.M., Dusterhoft A., Ehlich S.D., Emerson P.T.,
 Ertan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 Fritz K., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleon N.,
 Gilm S.Y., Glaeser B., Goffeau A., Golightly E.J., Grandi G.,
 Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones D.,
 Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,
 Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 Noone D., O'Reilly M., Ogawa K., Ogawara A., Oudega B., Park S.H.,
 Parro V., Pohl T.M., Portetelle D., Portolillo S., Prescott A.M.,
 Raeser M., Rajic P., Rumeille B., Rapoport G., Rey M., Reynolds S.,
 Plessem M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
 Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Solito B.,
 Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vaasaroti A.,
 Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,
 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 Yoshida K., Yoshikawa H.F., Zumberstein E., Yoshikawa H., Zanchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis";
 RL Nature 390:249-256 (1997).
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 malonyl-[acyl-carrier protein].
 CC -1- PATHWAY: Fatty acid biosynthesis.
 CC -1- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH
 TRANSACYLASE ACTIVITY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U59433; AAC44306.1; -;
 CC EMBL: Y13937; CA74249.1; -;
 CC EMBL: Z99112; CAB13463.1; -;
 CC HSSP: P25715; IMIA.

DR Subtilist; BG1836; fadb.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR004410; Fadb.
 DR Pfam; PF00598; Acyl_transf; 1.
 DR TIGRPFAMs; TIGR00126; fadb; 1.
 KW Fatty acid biosynthesis; transferase; Complete proteome.
 FT ACT SITE 91
 FT ACT SITE 91
 FT ACT SITE 201
 FT ACT SITE 201
 FT CONFLICT 124 131
 FT CONFLICT 124 131
 SQ SEQUENCE 317 AA; 34035 MW; 191AE828B1C9D17F CRC64;
 Query Match 16.9%; Score 269.5; DB 1; Length 317;
 Best Local Similarity 28.4%; Pred. No. 1.2e-14;
 Matches 81; Conservative 47; Mismatches 138; Indels 19; Gaps 8;
 QY 2 FVPPGGGAGWAGVRLVLLSSPVPFARMQACEALPFWDMWSVVDLLRDAGAVWERAD 61
 DB 6 FLFPGGGSGPTGKGLVEQVPARKLFDEADLTETKLSLIFE-----GDA--BELT 57
 QY 62 V--VQPVLFVWVSLAALMSYGIPEPDVLTGSHGEIAAAYGALSLDKAAVVALRS 118
 DB 58 LITNAGPALLTTSIVALEKFKESGITPDTAGSLGSEYSLVAGALSFPDVAVYTRKRG 117
 QY 119 RAL-AAVRGRGMAVSP-PIPAQVEQLIGE--RWAGRLVAAVNGPRSTANSGDAVND 173
 DB 118 EFWNEAVPAGEGMAAAILGMDAVALKQVTDKTEBGNLVQLANLNGQIVISGTAKGVE 177
 QY 174 EVLAACAGCVARRRIPVDYAS--HCPHVQPLREELLETGLDISPQSGVPFSTEGTW 231
 DB 178 LASELAKENAK-RALPLESGFSELMKPAKKEVADCDIDADVPIVSNVADY 236
 QY 232 LDTTTLDAVWYVNLHPVFPFSDVQALADGHRVFEVSPHPTL 276
 DB 237 MTEKADIKELIQLVSPVRFESINKLTAEGVTTETIEIGPKVL 281
 RESULT 13
 FADB SALTY STANDARD; PRT; 308 AA.
 ID FADB SALTY
 AC 085140;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Malonyl-CoA-acyl carrier protein transacylase (EC 2.3.1.39) (MCT).
 GN FADB OR STM1194.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2;
 RX MEDLINE=98317265; PubMed=9642179;
 RA Zhang Y., Cronan J.E. Jr.;
 RT "Transcriptional analysis of essential genes of the Escherichia coli
 fatty acid biosynthesis gene cluster by functional replacement with
 the analogous Salmonella typhimurium gene cluster";
 RL J. Bacteriol. 180:3295-3303 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SCS1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
 Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2";
 RL Nature 413:852-856 (2001).
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 malonyl-[acyl-carrier protein].
 CC -1- PATHWAY: Fatty acid biosynthesis.

-1- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH
TRANSLACYLASE ACTIVITY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcement/>
CC or send an email to license@ebi.ac.uk).

DR EMBL: AF044668; AAC38649.1; -
DR EMBL: AE008752; AAL20123.1; -
DR HSSP: P25715; 1MLA.
DR ScyGene: SG10637; fabd.
DR InterPro: IPR001227; Ac transferase.
DR InterPro: IPR004410; AcD.
DR Pfam: PF00698; Acyl_transfer_1.
DR TIGRPFam: TIGR00128; fabd; 1.
KM Fatty acid biosynthesis; Transferase; Complete proteome.
FT INIT MET 0 BY SIMILARITY.
FT ACT SITE 91 91 BY SIMILARITY.
FT ACT_SITE 200 200 BY SIMILARITY.
SQ SEQUENCE 308 AA; 32274 MW; B3565B45341A611A CRC64;

Query Match 16.0%; Score 256; DB 1; Length 308;
Best Local Similarity 28.9%; Pred. No. 1.4e-13;
Matches 87; Conservative 47; Mismatches 143; Indels 24; Gaps 11;

QY 2 FVPPGGGGQWAGNGVRLASSPVFARRMQACEBALPMDWSVVDILRDAGDAVERAD 61
DB 5 FVPPGGGGQWAGNGVRLASSPVFARRMQACEBALPMDWSVVDILRDAGDAVERAD 60
QY 62 VQVPLFVWVSLALMRSYGLPDAVIGHSGGEITAAHVCALSLKDAKTVVALRSRA 120
DB 61 -TQPALTLASVALMRWQGGGKMPALMAGHSAGEYSALVCAGINADRVLEVMKGR 119
QY 121 L--AAVRGRCM-ASVPLPAQVEQLIGERMAGRLMWAAV--NGPRSTAVSGDAEAVDEV 175
DB 120 MQDAVEEGTGMGAIIGLDASLAKACEBAGQV--VSPNPNFPGQVVAIGKKEAVERA 178
QY 176 LAYCACTGV-RARRIPVDVASHCPHVOPLREELLEIGDISPOSGVPFPSTVEGTWLD 234
DB 179 GAACKAGAKRALPLFVSPVSHCALMKPADKLAVELAKITFAPFVNVNVD---VK 234
QY 235 TTIIDAA----YWRNLHQPRPSDVAQALDDGHRVVEVSPHT---LYPAEDTTEDT 287
DB 235 CETDAALIRDALVRLVQVQWTKSVFIAAGVHELVGPGKVLGTLRIVDTLTAS 294
QY 288 A 288
DB 295 A 295

RESULT 14
FABD_ECOLI STANDARD; PRT; 308 AA.
ID _FABD_ECOLI
AC P25715;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39) (MCT).
GN FABD OR TFPA OR B1092.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
RC STRAIN=K12;
RX MEDLINE=92103950; PubMed=1339356;
RA Magnuson K., Oh W., Larson T.J., Cronan J.E. Jr.,
RT "Cloning and nucleotide sequence of the fabd gene encoding malonyl

RT coenzyme A-acyl carrier protein transacylase of Escherichia coli.";
RN FEBS Lett. 299:262-266 (1992).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92234941; PubMed=1314802;
RA Verweert I.I., Verdrete E.C., van der Linden K.H., Nijkamp H.J.,
RT Stultje A.R.;
RT "Cloning, nucleotide sequence, and expression of the Escherichia coli
RT fabd gene, encoding malonyl coenzyme A-acyl carrier protein
RT transacylase";
RL J. Bacteriol. 174:2851-2857 (1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95214624; PubMed=7700236;
RA Bouguin N., Tempete M., Holland I.B., Seror S.J.;
RT "Resistance to trifluoroperazine, a calmodulin inhibitor, maps to the
RT fabd locus in Escherichia coli";
RL Mol. Gen. Genet. 246:628-637 (1995).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick J.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1233-1238 (1997).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Ikeda T., Itoh T., Kajiya M., Kanai K., Kashimoto K.,
RA Kimura S., Kitegawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horikuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map";
RL DNA Res. 3:137-155 (1996).
RN [6]
RP SEQUENCE OF 288-308 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92210530; PubMed=1556094;
RA Cronan J.E. Jr., Rawlings M.;
RT "The gene encoding Escherichia coli acyl carrier protein lies within
RT a cluster of fatty acid biosynthetic genes";
RL J. Biol. Chem. 267:5751-5754 (1992).
RN [7]
RP SEQUENCE OF 1-11.
RC STRAIN=K12 / W3110;
RA Pasquall C., Sanchez J.-C., Ravler F., Golaz O., Hughes G.J.,
RA Fruhlinger S., Paquet N., Wilkins M., Appel R.D., Bairoch A.,
RA Hochstrasser D.F.;
RT Submitted (SEP-1994) to the SWISS-PROT data bank.
RN [8]
RP SEQUENCE OF 1-10.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12";
RL Electrophoresis 18:1259-1313 (1997).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX MEDLINE=95286570; PubMed=7768883;
RA Serre L., Verdrete E.C., Dautez Z., Stultje A.R., Derewenda Z.S.;
RT "The Escherichia coli malonyl-CoA:acyl carrier protein transacylase
RT at 1.5-A resolution. Crystal structure of a fatty acid synthase
RT component";
RL J. Biol. Chem. 270:12961-12964 (1995).
CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +

CC malonyl-[acyl-carrier protein].
 CC -1- PATHWAY: Fatty acid biosynthesis.
 CC -1- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH
 CC TRANSACYLASE ACTIVITY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M87040; AAA23742.1; -
 DR EMBL; Z11565; CAAT7658.1; -
 DR EMBL; M84991; AAA23738.1; -
 DR EMBL; AB000210; AAC74176.1; -
 DR EMBL; D90745; BAA35900.1; -
 DR PIR; S20443; S20443.
 DR PIR; B41856; B41856.
 DR PDB; 1MLA; 25-JAN-95.
 DR SWISS-2DPAGE; P25715; COLI.
 DR EcGene; EG1317; fabD.
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR004410; FabD.
 DR Pfam; PF00698; Acyl transferase, 1.
 DR TIGRFAMs; TIGR00126; fabD; 1.
 KM Fatty acid biosynthesis; Transferase; 3D-structure; Complete proteome.
 FT INIT MET 0
 FT ACT_SITE 91 91
 FT ACT_SITE 200 200
 SQ SEQUENCE 308 AA; 32286 MW; F5901043D92PED8 CRC64;
 Query Match 15.9%; Score 254; DB 1; Length 308;
 Best Local Similarity 29.1%; Pred. No. 2,1e-13; Indels 20; Gaps 11;
 Matches 87; Conservative 47; Mismatches 145;
 QY 2 FVPPGQGGQWAGMGVRLAASSPVFARMQACEALAWVWVSVDILRRDAGDAVNERAD 61
 DB 5 FVPPGQGGQWAGMGVRLAASSPVFARMQACEALAWVWVSVDILRRDAGDAVNERAD 61
 QY 62 VVQPVLFVWVSLAALRSYVGIKEDAVL-GHSGCEIAAHVCGALSKDAKTYALRSRA 120
 DB 61 -TPPALTATVAVLRVWQGGKAPAMWAGSHLGEYSALVACAVIDPADVRLVEMRGKE 119
 QY 121 L--AAVARGGMA--VLPAPQEVNQILGERMAGLWAAV--NPPRTAVSGDAEAVDEV 175
 DB 120 MQSLVPEGTAMMAITGLDPASTAKACEAEAGV-VSPVNFNSPGQVVIAGHKEAVERA 178
 QY 176 LAYCAGTGV-RARRIPDVASHCHPVOPLREELLELLGDISPPQSPGVFPFSTVEGTWLD 234
 DB 179 GAACKAGARALPLPVSPVSHCLMKMRPADKLAVALAKITFNAPTYPVVANNVDVK--E 236
 QY 235 TTLDAA--WYRNLIHQVRFPSDAVQALDDGHRVFEVSPHPT--LVPAIEDTTEDTA 288
 DB 237 TNGDAIRDALVRQLNPNVQWTKSVYMAAGVEHLVEVGPKVLTGLTKETVDTLTASA 295
 RESULT 15
 ID FAS_RAT STANDARD; PRT; 2505 AA.
 AC P12785; O64717; O09187; O09190;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [includes: EC 2.3.1.38; EC 2.3.1.39;
 DE EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=69240686; PubMed=2717611;
 RA Amy C.M., Witkowski A., Naggett J., Williams B., Randhawa Z.,
 RA Smith S.;
 RT "Molecular cloning and sequencing of cDNAs encoding the entire rat
 RT fatty acid synthase."
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3114-3118(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=93075999; PubMed=1339331;
 RA Beck K.F., Schreglmann R., Stathopoulos I., Klein H., Hoch J.,
 RA Schweitzer M.;
 RT "The fatty acid synthase (FAS) gene and its promoter in Rattus
 RT norvegicus."
 RL DNA Seq. 2:359-386(1992).
 RN [3]
 RP SEQUENCE OF 75-2505 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Mammary gland;
 RX MEDLINE=69128431; PubMed=2915923;
 RA Schweitzer M., Takabayashi K., Beck K.F., Schreglmann R.;
 RT "rat mammary gland fatty acid synthase: localization of the
 RT constituent domains and two functional polyadenylation/termination
 RT signals in the cDNA."
 RL Nucleic Acids Res. 17:567-586(1989).
 RN [4]
 RP SEQUENCE OF 2085-2505 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=88087240; PubMed=2891707;
 RA Naggett J., Witkowski A., Mikkelsen J., Smith S.;
 RT "Molecular cloning and sequencing of a cDNA encoding the thioesterase
 RT domain of the rat fatty acid synthetase."
 RL J. Biol. Chem. 263:1146-1150(1988).
 RN [5]
 RP SEQUENCE OF 1921-2324 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=87246646; PubMed=3109907;
 RA Witkowski A., Naggett J., Mikkelsen J., Smith S.;
 RT "Molecular cloning and sequencing of a cDNA encoding the acyl carrier
 RT protein and its flanking domains in the mammalian fatty acid
 RT synthetase."
 RL Eur. J. Biochem. 165:601-606(1987).
 CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROTEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) CoA + N CO(2) + 2N NADP(+).
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxypalmitoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-
 CC carrier protein] + oleate.
 CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL, M76767, AAA57219.1, -
DR EMBL, X62888, CAA44679.1, -
DR EMBL, X62889, CAA44680.1, -
DR EMBL, X13415, CAA31780.1, -
DR EMBL, X13527, CAA31882.1, -
DR EMBL, U03514, AAA41144.1, -
DR PIR, A30313, XYRTFA.
DR InterPro, IPR001227, AC transferase.
DR InterPro, IPR002085, Adh zn family.
DR InterPro, IPR000794, ketoacyl-synt.
DR InterPro, IPR003880, pantone attach.
DR InterPro, IPR000051, SAM_bind.
DR InterPro, IPR001031, Thioesterase.
DR Pfam, PF00107, adh_zinc, 1.
DR Pfam, PF00109, ketoacyl-synt, 1.
DR Pfam, PF00550, pp-binding, 1.
DR Pfam, PF00598, acyl_transfer, 1.
DR Pfam, PF00975, Thioesterase, 1.
DR Pfam, PF02801, ketoacyl-synt_C, 1.
DR PROSITE, PS00012, PHOSPHOPANTETHEINE, 1.
DR PROSITE, PS00606, B KETOACYL SYNTHASE, 1.
DR PROSITE, PS50075, ACP DOMAIN, 1.
KM Fatty acid biosynthesis, Multifunctional enzyme; Phosphopantetheine;
KM Hydrolyase; Oxidoreductase; transferase; Lyase; NADP;
KW Pyridoxal phosphate.
FT DOMAIN 1 413 BETA-KETOACYL SYNTHASE.
FT DOMAIN 429 817 ACYL AND MALONYL TRANSFERASES.
FT DOMAIN 1629 1857 ENOYL REDUCTASE.
FT DOMAIN 1858 2113 BETA-KETOACYL REDUCTASE.
FT DOMAIN 2118 2174 ACYL CARRIER (ACP).
FT DOMAIN 2202 2505 THIOESTERASE.
FT ACT SITE 161 161 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT ACT SITE 581 581 MALONYLTRANSFERASE (BY SIMILARITY).
FT NP_BIND 1665 1682 NADP (ER).
FT BINDING 1698 1698 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT NP_BIND 1765 1780 NADP (KR).
FT BINDING 2151 2151 PHOSPHOPANTETHEINE (BY SIMILARITY).
FT ACT SITE 2302 2302 THIOESTERASE (BY SIMILARITY).
FT ACT SITE 2475 2475 THIOESTERASE (BY SIMILARITY).
FT ACT_SITE 878 878 BETA-HYDROXYACYL DEHYDRATASE (BY SIMILARITY).
FT CONFLICT 871 871 S -> P (IN REF. 3).
FT CONFLICT 1967 1968 MV -> IL (IN REF. 5).
FT CONFLICT 2085 2085 C -> P (IN REF. 4).
FT CONFLICT 2106 2106 A -> V (IN REF. 1 AND 5).
FT CONFLICT 2296 2296 Y -> H (IN REF. 1 AND 5).
SQ SEQUENCE 2505 AA, 272647 MW, 5810EC13D37F3114 CRC64;

Query Match 15.24; Score 243; DB 1; Length 2505;
Best Local Similarity 27.54; Pred. No. 1.9e-11;
Matches 87; Conservative 56; Mismatches 149; Indels 24; Gaps 13;

QY 2 FVPPGGGQWAGNGVRLASSPVFARRMQACEALAPWDSVVDIIRRDAGDAWERAD 61
DB 494 FICSGMGTQWRGGLSLMRDS-FRESILRSDEALKP-LGVKVSDDL-LSTDEHTFD-D 548
QY 62 VVQPV--LPSVWVSLAAMRSYGIERDAVVGSGGEIAAAVCGALSLLKDAKTVALRSR 119
DB 549 IVHSFVSLRIQIALDLITSMGLKPDGIIHSLGEVAGIADGCLSGREAVLAAYWRQ 608
QY 120 ALA-AVRGRGMAVPLPAQVEVEQLIGRWAGRLWVAVNGPRSTAVSGDAEAVDEVLAY 178
DB 609 CIDANLPAGSMNAVGSWECKQ---RCPGVVPACHNSEDTVTISGPAVNEFVEQ 664
QY 179 CAGTGVARIRIV-DVASHCPHVQPLRELLLELGLDI--SPQSGVPFFST--VEGTWLD 233
DB 665 LKQEGVFAKEVRIGLAFHSYFMEGIALPTLQALKVIRPRPRASRWLSTSIPEAQMS 724
QY 234 T--TTLDAAAYVYRNLIQPVPSDAVQALDDGHRVPEVSPHPTLVPAIEDTETDAEDV 291
DB 725 SLARTSSAEYVNNVNLSPVLFQELHWPE--HAVVEIAPHALQAVLKRGVKPSG--- 779
QY 292 TAIGSLRGDNDTRRF 307

Db 780 TIPLMKRDHKDNIIEF 795

Search completed: June 17, 2003, 13:02:34
Job time : 5.34793 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 17, 2003, 12:55:32 ; Search time 6.85529 Seconds
(without alignments)
4305.183 Million cell updates/sec

Title: US-09-914-286-3_COPY_1050_1356

Perfect score: 1598

Sequence: 1 VVFVPCQCGQGMAGMGRVLLA.....AEDEVTAIGSLRQDNDTRRF 307

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Selecting first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	879.5	55.0	3573	2 S23070	erythronolide synt
2	858	53.7	3519	2 S43048	polyketide synthas
3	848.5	53.1	5069	2 T17464	rifamycin polyketi
4	840.5	52.6	3491	2 T43231	probable 6-deoxyer
5	828.5	51.8	2124	2 T28658	polyketide synthas
6	815.5	51.0	4735	2 T17463	rifamycin polyketi
7	812.5	50.8	1763	2 T17465	rifamycin polyketi
8	807	50.5	7576	2 T17428	FK506 polyketide s
9	806.5	50.5	3413	2 T17467	rifamycin polyketi
10	805	50.4	1728	2 T17466	rifamycin polyketi
11	790.5	49.5	2126	2 H70621	probable polyketid
12	790	49.4	3172	2 S22012	erythronolide synt
13	790	49.4	3178	2 S13595	6-deoxyerythronol
14	784	49.1	4151	2 G70944	probable polyketid
15	775.5	48.5	1937	2 T03224	probable polyketid
16	772.5	48.3	1346	2 T17412	polyketide synthas
17	772.5	48.3	3739	2 T17410	polyketide synthas
18	772	48.3	6260	2 T30228	polyketide synthas
19	772	48.3	10233	2 T30225	polyketide synthas
20	769	48.1	4613	2 T17409	polyketide synthas
21	766.5	48.0	8563	2 T30226	polyketide synthas
22	763	47.7	1562	2 T17411	polyketide synthas
23	749	46.7	2723	2 T03221	probable polyketid
24	747	46.7	2100	2 T03223	probable polyketid
25	736	46.0	6420	2 T30283	polyketide synthas
26	575.5	35.6	2108	2 H70819	probable polyketid
27	570	35.7	2110	2 B44110	mycocerosate synth
28	570	35.7	2111	2 A70668	mycocerosate synth
29	561.5	35.1	1582	2 E70876	probable polyketid

30	561.5	35.1	2118	2 S72705	mycocerosate synth
31	559.5	35.0	1570	2 AC2012	hypothetical prote
32	559	35.0	2116	2 C86926	probable mycoceros
33	542.5	33.9	1293	2 T30871	orsellinic acid sy
34	532.5	33.3	1837	2 B70984	probable polyketid
35	530.5	33.2	2126	2 E70822	probable polyketid
36	500	31.3	635	2 S73017	polyketide synthas
37	500	31.3	1446	2 S73013	polyketide synthas
38	500	31.3	1540	2 H87203	polyketide synthas
39	500	31.3	1871	2 A87204	polyketide synthas
40	493.5	30.9	1815	2 S73021	polyketide synthas
41	493.5	30.9	1822	2 F87203	polyketide synthas
42	493	30.9	1538	2 E70874	probable psb prot
43	483	30.2	2188	2 A70984	probable polyketid
44	472.5	29.6	497	2 S37580	probable acyltrans
45	472.5	29.6	497	2 S37581	probable acyltrans

ALIGNMENTS

RESULT 1

S23070 erythronolide synthase (EC 2.3.1.94) II - Saccharopolyspora erythraea

N/Alternate names: 6-deoxyerythronolide B synthase II

C/Species: Saccharopolyspora erythraea

C/Date: 07-Apr-1994 #sequence revision 07-Apr-1994 #ext_change 21-Jul-2000

C/Accession: S23070; S22011; S23205

R/Bevitt, D.J.; Cortes, J.; Haydock, S.F.; Leadlay, P.F.

Eur. J. Biochem. 204, 39-49, 1992

A/Title: 6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea. Cloning of

A/Reference number: S23070; WUID:92155230; PMID:1740151

A/Accession: S23070

A/Molecule type: DNA

A/Residues: 1-3573 <BEV1>

A/Cross-references: EMBL:X62569

A/Experimental source: strain NRRL 2338

R/Bevitt, D.J.

submitted to the EMBL Data Library, September 1991

A/Reference number: S22011

A/Accession: S22011

A/Molecule type: DNA

A/Residues: 1-184, 'I', 186-301, 'S', 303-521, 523-658, 'A', 660-993, 1001-1212, 'H', 1214-1392, 1

3479, 'DH', 3480-3572 <BEV2>

A/Cross-references: EMBL:X62569; NID:946977; PIDN:CA44448.1; PID:9581651

R/Gaffney, P.; Bevitt, D.J.; Staunton, J.; Leadlay, P.F.

FEBS Lett. 304, 225-228, 1992

A/Title: Identification of DEBS 1, DEBS 2 and DEBS 3, the multienzyme polypeptides of t

A/Reference number: S23103; WUID:92316235; PMID:1618327

A/Accession: S23205

A/Molecule type: protein

A/Residues: 2-12, 'XXX' <CAF>

A/Experimental source: strain CA340

C/Genetics:

A/Genes: eryA

A/Start codon: GNG

C/Function:

A/Description: catalyzes the construction of a polyketide chain, which is then cyclised

A/Pathway: erythromycin biosynthesis

C/Superfamily: [acyl-carrier-protein] S-malonyltransferase homology; 3-oxoacyl-[acyl-ca

riochain alcohol dehydrogenase homology

C/Keywords: acyltransferase; antibiotic biosynthesis; carrier protein; coenzyme A; mult

F:561-843/Domain: [acyl-carrier-protein] synthase I homology <OAS1>

F:1140-1308/Domain: short-chain alcohol dehydrogenase homology <AWT1>

F:1404-1475/Domain: acyl carrier protein homology <ACP1>

F:1519-1919/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>

F:2023-2305/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AWT2>

F:2857-3131/Domain: long-chain alcohol dehydrogenase homology <LADH>

F:3149-3337/Domain: short-chain alcohol dehydrogenase homology <SADH>

F:3420-3493/Domain: acyl carrier protein homology <ACP2>

F:1439/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

[illegible]

RESULT 7
T17465
rifamycin polyketide synthase module 7 - Amycolatopsis mediterranei
C:Species: Amycolatopsis mediterranei
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 03-Nov-2000
C:Accession: T17465
R:Schupp, T.
submitted to the EMBL Data Library, December 1997
A:Reference number: Z18802
A:Accession: T17465
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1763 <SCH>
A:Cross-references: EMBL/AJ223012; NID:e1227119; PID:e1227122; PIDN:CAA11037.1
A:Experimental source: strain IBG A3136
C:Superfamily: Streptomyces hygroscopicus probable polyketide synthase module 4; 3-oxoad
homology; [acyl-carrier-protein] S-malonyltransferase homology
C:Keywords: carrier protein
/:1615-1686/Domain: acyl carrier protein homology <ACP>

Query Match	50.8%	Score	812.5	DB	2	Length	1763
Best Local Similarity	55.2%	Pred. No.	7.6e-55				
Matches	170	Conservative	34	Mismatches	93	Indels	11
						Gaps	5

QY	1	VFEVFGGCGGAGMGVRLTASSPFARMOQACEEALAWVMWSVVDILIRPDAGAGVWMERRA	60
Db	555	VWVFPGCGTQGVMGKRELLDASPVFAERIKEGCAALDQMTWSLLDVLVR---	GGGDDLSV 611
QY	61	DVVOQPVLFVWVSLAALMRSYGIREDVATLGHSGGEBIAAAHYCGALSTLKDAAKTVALKSRA	120
Db	612	EVLQAPACAVWVWGLAAVWMSGVGRDVAVHGSGGBIAAACVSGALTLLDAAKVALRSQA	671
QY	121	LAA-VRGGGMAVPLPAQOEVEOLIGBRMAGSLWAAVNGPRSTVSQDAEAVENVLAYC	179
Db	672	IAAHLISGGGMAVSLSEDEANNRIG-LMDGRILEVAANGNPASVVIADDAQALDEALEVL	730
QY	180	AGTGVRAIRIPVDYASHCPHVQPLREELLELLIGDISIPQSGVPEPSTVEGFWL-DTTTLD	238
Db	731	AGDGVRAVQVADVASHTRHVEDIRDTLAETLAGITTAQAPDVPFESTVTGGMVADADVL	790
QY	239	AAWYRNRLHQVRFEDVAQALADDOHRVFWVSPPHPTVPLAIEDTETDJAADVYALISLR	298
Db	791	GGYWRNRRLNQRREFPAVAELLEQGHGVFEVASPAVLEVQPSIELT-----DAVVTGLR	845

Qy	299	RGDNDTRR	306
Db	846	RDDGGLRR	853

RESULT 8
T17428
FK506 polyketide synthase - Streptomyces sp. (strain MA6548)
C:Species: Streptomyces sp.
A:Variety: strain MA6548
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000
Accession: T17428

R:Motamedi, H.,Shafiee, A.
Eur. J. Biochem. 256: 528-534, 1998
A>Title: The biosynthetic gene cluster for the macrolactone ring of the immunosuppressor
A:Reference number: Z18779; MUID:98451508; PMID:9780228
A:Accession: J17428
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7576 <NOT>
A:Cross-references: EMBL:AF082100; NID:g3798623; PID:g3798624; PIDN:ACC68815.1
A:Experimental source: Strain MA6548
C:Genetics:
A:Gene: fkbB

A>Description: involved in synthesis of the backbone of the immunosuppressant FK506 poly C/Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acetate-CoA ligase C/Keywords: carrier protein

F.154-500/Domain: acetate-CoA ligase homology <ACL>
F.1095-1166/Domain: acyl carrier protein homology <ACP1>
F.11204-1599/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F.11662-1953/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F.12680-2751/Domain: acyl carrier protein homology <AC2>
F.12804-3198/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F.13285-3569/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F.14330-4391/Domain: acyl carrier protein homology <AC3>
F.14435-4830/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F.15903-5974/Domain: acyl carrier protein homology <ACP4>
F.6018-6412/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F.6513-6785/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
F.7450-7521/Domain: acyl carrier protein homology <AC5>

Query Match	50.5%	Score 807	DB 2	Length 7576
Best Local Similarity	55.7%	Pred. No. 1.2e-53		
Matches 1/72	Conservative 34	Mismatches 89	Indels 14	Gaps 7

```

QY 1 VFVPRGGCGAGNAGVNTLLASPFPARBMQCEALAPMVDMSVYDLIRGACGAVMBERA 60
Db 3295 VFVFPGGQSQWDGNGAELLATEPVPFARLIGCAALALYTWMDLLDVARRPGAEIDRV 3354
QY 61 DVQVPLVFSVWVSLAALMRSYGIPEPDVILGHSQGEIAAAHYCGALSLDKDAKTALRSRA 120
Db 3355 DVQVQASAPMMVALAELMRAGVAPAAVVGHSQGEVAAACVAGTUTLDDAAKVALRSRL 3414
QY 121 LAAVR-GRGEMASVLPRAOEVEOLIGERWAGRLVAAVNGBRSTAVSGDAEAVDEVLYC 179
Db 3415 VATRAGHGGAVSWVP-PA----DDAAMAAGRLEVAANVGPASTLVAGAADAVEELLA-- 3464
QY 180 AGTVRARRIVDYASHCHQVPIRBEELIELGDISPOQSGVPPSPSTYEGVWLDITTLIDA 239
Db 3468 --ATPHARRIADVASHTAHVESIRGALLDLADLTGPAEPIPEFSTVDEAMLRPA--DA 3524
QY 240 AYWVRNLHQVPRFSDAVQALDDHRAVVEVSPHPTLVPALEDITTEDVAAEAVTDI--GSLR 298
Db 3525 AYWVDNVCAPRFGAALAAARLAEIQRHVRFEVSPHPVLTTLADTL--AGHPNTAVTGTLR 3584
QY 299 RGDNDRTRF 307
Db 3583 RGDGGARRF 3591

```

RESULT 9

C:Species: Amycolatopsis mediterranei
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
C:Accession: T17467

submitted to the EMBL Data Library, December 1997
A:Reference number: Z18802
A:Accession: T17467
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3413 <SCH>
A:Cross-references: EMBL:AJ223012; NID:e1227119; PID:e1227124; PIDD:CAA11039.1

C/Species: Saccharopolyspora erythraea
 C/Date: 07-Apr-1994 #sequence revision 07-Apr-1994 #text_change 26-May-2000
 C/Accession: S22012, S34784, S23206
 R/Bevitt, D.J.
 Submitted to the EMBL Data Library, September 1991
 A/Reference number: S22011
 A/Accession: S22012
 A/Molecule type: DNA
 A/Residues: 1-3172 <BEV1>
 A/Cross-references: EMBL:X62569; NID:g46977; PIRN:CAA44449.1; PID:g46979
 A/Experimental source: strain NRRL 2338
 R/Bevitt, D.J.; Cortes, J.; Haydock, S.F.; Leadlay, P.F.
 Eur. J. Biochem. 204, 39-49, 1992
 A/Title: 6-Deoxyerythronolide B synthase 2 from Saccharopolyspora erythraea. Cloning of
 A/Reference number: S23070; MUID:92155230; PMID:11740151
 A/Accession: S34784
 A/Molecule type: DNA
 A/Residues: 1-715, 717-728, 'C', 729-886, 888-894, 'QR', 895-910, 'E', 912-1276, 'RR', 1279-1284, 'EV2'
 A/Cross-references: EMBL:X62569
 R/Caffrey, P.; Bevitt, D.J.; Staunton, J.; Leadlay, P.F.
 FEBS Lett. 304, 225-228, 1992
 A/Title: Identification of DEBS 1, DEBS 2 and DEBS 3, the multienzyme polypeptides of the
 A/Reference number: S23103; MUID:92316235; PMID:1618327
 A/Accession: S23206
 A/Molecule type: protein
 A/Residues: 2-11, 'XXX', 15 <CAF>
 A/Experimental source: strain CN340
 C/Genetics:
 A/Gene: eryA
 C/Function:
 A/Description: catalyzes the construction of a polyketide chain, which is then cyclised
 C/Superfamily: [acyl-carrier-protein] S-malonyltransferase homology; 3-oxoacyl-[acyl-carrier-protein] short-chain alcohol dehydrogenase homology
 C/Keywords: acyltransferase; antibiotic biosynthesis; carrier protein; coenzyme A; multi-
 F/60-450/Domain: [acyl-carrier-protein] synthase I homology <OAS1>
 F/555-833/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
 F/1117-1297/Domain: short-chain alcohol dehydrogenase homology <SAD1>
 F/1192-1463/Domain: acyl carrier protein homology <ACP1>
 F/1511-1914/Domain: [acyl-carrier-protein] synthase I homology <OAS2>
 F/2022-2300/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
 F/2557-2733/Domain: short-chain alcohol dehydrogenase homology <SAD2>
 F/2819-2896/Domain: acyl carrier protein homology <ACP2>
 F/2959-3153/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>
 Query Match 49.4%; Score 790; DB 2; Length 3172;
 Best Local Similarity 56.0%; Pred. No. 8.6e-53;
 Matches 169; Conservative 39; Mismatches 84; Indels 10; Gaps 5;
 Db 3 VFPQGGQWAGMVGVRLLASSPVFARMQACEALAPWDMVSVDILRRDAGDAVBERADV 62
 557 VFPQGGQWAGMVGVRLLASSPVFARMQACEALAPWDMVSVDILRRDAGDAVBERADV 613
 Qy 63 VQPLFVWVSLAALMRSYGIEPDVAVLGHSGEITAAHVAGALSLDKDAKTVALLRSALA 122
 Db 614 VQPLFVWVSLAALMRSYGIEPDVAVLGHSGEITAAHVAGALSLDKDAKTVALLRSALA 673
 Qy 123 AVRGQGMASVPLPAQVEBQILGERMAGRLVAAVNGPRS-TAVSGDAEAVDEVLAYCAG 181
 Db 674 RLGGQGMASVPLPAQVEBQILGERMAGRLVAAVNGPRS-TAVSGDAEAVDEVLAYCAG 732
 Qy 182 TGVARRIPVDYASHCHQVOPLEELLEGISPPSGVFPFSTVGTMLDITTDAA 241
 Db 733 EAHKARRIPVDYASHCHQVOPLEELLEGISPPSGVFPFSTVGTMLDITTDAA 792
 Qy 242 WYRNLLQPVFSDAVOALDGHVFEVSEPHPTLVPAIEDTETDAEDVTA-----IGSL 297
 Db 793 WYRNLLQPVFSDAVOALDGHVFEVSEPHPTLVPAIEDTETDAEDVTA-----IGSL 851
 Qy 298 RR 299
 Db 852 RR 853

RESULT 13
 S13595
 6-deoxyerythronolide B synthase - Saccharopolyspora erythraea
 N/Alternate names: protein A
 C/Species: Saccharopolyspora erythraea
 C/Date: 07-Apr-1994 #sequence revision 07-Apr-1994 #text_change 26-May-2000
 C/Accession: S13595
 R/Cortes, J.; Haydock, S.F.; Roberts, G.A.; Bevitt, D.J.; Leadlay, P.F.
 Nature 348, 176-178, 1990
 A/Title: An unusually large multifunctional polypeptide in the erythromycin-producing pc
 A/Reference number: S13595; MUID:91043075; PMID:2234082
 A/Accession: S13595
 A/Molecule type: DNA
 A/Residues: 1-3178 <COR>
 A/Cross-references: EMBL:X6107; NID:g46975; PIRN:CAA39583.1; PID:g46976
 A/Experimental source: strain NRRL 2338
 C/Genetics:
 A/Gene: eryA
 C/Function:
 A/Description: catalyzes the construction of a polyketide chain, which is then cyclised
 A/Pathway: erythromycin biosynthesis
 C/Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protein
 C/Keywords: acyltransferase; antibiotic biosynthesis; carrier protein; multifunctional
 F/60-450/Domain: [acyl-carrier-protein] S-malonyltransferase homology
 F/553-833/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
 F/1123-1303/Domain: short-chain alcohol dehydrogenase homology <SAD1>
 F/1398-1469/Domain: acyl carrier protein homology <ACP1>
 F/1517-1920/Domain: [acyl-carrier-protein] synthase I homology <OAS2>
 F/2028-2300/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
 F/2563-2740/Domain: short-chain alcohol dehydrogenase homology <SAD2>
 F/2825-2896/Domain: acyl carrier protein homology <ACP2>
 F/2965-3158/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>
 Query Match 49.4%; Score 790; DB 2; Length 3178;
 Best Local Similarity 56.0%; Pred. No. 8.6e-53;
 Matches 169; Conservative 39; Mismatches 84; Indels 10; Gaps 5;
 Db 3 VFPQGGQWAGMVGVRLLASSPVFARMQACEALAPWDMVSVDILRRDAGDAVBERADV 62
 555 VFPQGGQWAGMVGVRLLASSPVFARMQACEALAPWDMVSVDILRRDAGDAVBERADV 611
 Qy 63 VQPLFVWVSLAALMRSYGIEPDVAVLGHSGEITAAHVAGALSLDKDAKTVALLRSALA 122
 Db 612 VQPLFVWVSLAALMRSYGIEPDVAVLGHSGEITAAHVAGALSLDKDAKTVALLRSALA 671
 Qy 123 AVRGQGMASVPLPAQVEBQILGERMAGRLVAAVNGPRS-TAVSGDAEAVDEVLAYCAG 181
 Db 672 RLGGQGMASVPLPAQVEBQILGERMAGRLVAAVNGPRS-TAVSGDAEAVDEVLAYCAG 730
 Qy 182 TGVARRIPVDYASHCHQVOPLEELLEGISPPSGVFPFSTVGTMLDITTDAA 241
 Db 731 EAHKARRIPVDYASHCHQVOPLEELLEGISPPSGVFPFSTVGTMLDITTDAA 790
 Qy 242 WYRNLLQPVFSDAVOALDGHVFEVSEPHPTLVPAIEDTETDAEDVTA-----IGSL 297
 Db 791 WYRNLLQPVFSDAVOALDGHVFEVSEPHPTLVPAIEDTETDAEDVTA-----IGSL 849
 Qy 298 RR 299
 Db 850 RR 851

RESULT 14
 G70944
 probable polyketide synthase Rv2048c - Mycobacterium tuberculosis (strain H37RV)
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence revision 17-Jul-1998 #text_change 20-Jun-2000
 C/Accession: G70944
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.

```

F:55-456/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F:549-827/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>

Query Match          48.5%; Score 775.5; DB 2; Length 1937;
Best Local Similarity 53.1%; Pred. No. 6,4e-52;
Matches 161; Conservative 37; Mismatches 96; Indels 9; Gaps 3;

QY 1 VFVPRGGCGGMAGGVULLASSPPVFARRMQACEBALAPMYDMSVVDILRRDAGDAVNERA 60
   |||||
DB 549 VFVPRGGGGSQWIGMGRGLMESSAVFRDSVLRTRDRLAEFVDWSMAAVALAGEPGTPDLDRV 608
   |||||

QY 61 DVVQPVLFVYVNSLAALMRSYGIEPDAVLGHSGEIAAAHYCGALSLKDAKTVALLRSRA 120
   |||||
DB 609 DVVQPVLFYVYVNSLAALMRSYGIEPDAVLGHSGEIAAAHYCGALSLKDAKTVALLRSRA 668
   |||||

QY 121 LAAVVRGGRGMAASVPLPAQEEVEOLIGERRMAGRLWVAANVGPRSTAVSGDAEAVDEVLAYCA 180
   |||||
DB 669 LTLALAGEGGMVWVQRSVAAYVENLL-RQWEGRLSYAAVVGPEAVVSGQVALBELLA--- 724
   |||||

QY 181 GTGVRARRIRPDYASHCPHYQPLREELLELGLDISPQSGVPPFPSTEGTMDLDTTTIDAA 240
   |||||
DB 725 -TEDDRARRVAVDYASHSAQVERIEEKLRTLTLDVQPMTSRVPLEFSTVERDMIDTASMDTG 783
   |||||

QY 241 YWYRNLLHQVRFPSDAVQALADDGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 300
   |||||
DB 784 YWYRNLLHQVRFPSDAVQALADDGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 839
   |||||

QY 301 DND 303
   ||
DB 840 HDD 842

Search completed: June 17, 2003, 13:12:21
Job time : 8.85529 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using BW model

Run on: June 17, 2003, 13:08:02 ; Search time 9.56552 Seconds

(without alignments)
3431.399 Million cell updates/sec

Title: US-09-914-286-3_COPY_1050_1356

Perfect score: 1598
Sequence: 1 VFVFPQGGQGMAGVRLA.....AEDVTAIGLRGNDNTRRP 307

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB pep:*
2: /cgn2_6/prodata/1/pubpaa/PCR_NEW_PUB pep:*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB pep:*
6: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB pep:*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB pep:*
10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	891.5	55.8	3816	US-09-808-880-3	Sequence 3, Appl1
2	861.5	53.9	4150	US-09-808-880-2	Sequence 2, Appl1
3	858	53.7	3519	US-09-808-880-4	Sequence 4, Appl1
4	828	51.8	344	US-09-735-056-33	Sequence 33, Appl1
5	772.5	48.3	1346	US-09-860-846-37	Sequence 37, Appl1
6	772.5	48.3	1346	US-09-988-3848-37	Sequence 37, Appl1
7	772.5	48.3	1346	US-09-836-821-37	Sequence 37, Appl1
8	772.5	48.3	1346	US-09-793-708-4	Sequence 4, Appl1
9	772.5	48.3	1346	US-09-861-289-37	Sequence 37, Appl1
10	772.5	48.3	3739	US-09-860-846-33	Sequence 33, Appl1
11	772.5	48.3	3739	US-09-988-3848-33	Sequence 33, Appl1
12	772.5	48.3	3739	US-09-836-821-33	Sequence 33, Appl1
13	772.5	48.3	3739	US-09-793-708-2	Sequence 2, Appl1
14	772.5	48.3	3739	US-09-861-289-33	Sequence 33, Appl1
15	772.5	48.3	11877	US-09-860-846-6	Sequence 6, Appl1
16	772.5	48.3	11877	US-09-836-821-6	Sequence 6, Appl1
17	772.5	48.3	11877	US-09-861-289-6	Sequence 6, Appl1
18	772.5	48.3	12199	US-09-988-3848-6	Sequence 6, Appl1
19	772	48.3	345	US-09-735-056-34	Sequence 34, Appl1

20	769	48.1	4551	US-09-793-708-1	Sequence 1, Appl1
21	769	48.1	4613	US-09-860-846-31	Sequence 31, Appl1
22	769	48.1	4613	US-09-988-3848-31	Sequence 31, Appl1
23	769	48.1	4613	US-09-836-821-31	Sequence 31, Appl1
24	769	48.1	4613	US-09-861-289-31	Sequence 31, Appl1
25	763	47.7	1562	US-09-860-846-35	Sequence 35, Appl1
26	763	47.7	1562	US-09-988-3848-35	Sequence 35, Appl1
27	763	47.7	1562	US-09-836-821-35	Sequence 35, Appl1
28	763	47.7	1562	US-09-793-708-3	Sequence 3, Appl1
29	763	47.7	1562	US-09-861-289-35	Sequence 35, Appl1
30	732.5	45.8	7257	US-10-014-717-5	Sequence 5, Appl1
31	723.5	45.3	1832	US-10-014-717-4	Sequence 4, Appl1
32	723.5	45.3	3798	US-10-014-717-6	Sequence 6, Appl1
33	532.5	33.3	1827	US-09-712-363-261	Sequence 261, App
34	490	30.7	5215	US-09-860-846-2	Sequence 2, Appl1
35	490	30.7	5215	US-09-988-3848-2	Sequence 2, Appl1
36	490	30.7	5215	US-09-836-821-2	Sequence 2, Appl1
37	490	30.7	5215	US-09-861-289-2	Sequence 2, Appl1
38	440.5	27.6	1402	US-09-712-363-166	Sequence 166, App
39	428.5	26.8	343	US-09-735-056-32	Sequence 32, App
40	413.5	25.9	1616	US-09-712-363-262	Sequence 262, App
41	400.5	25.1	2439	US-10-014-717-7	Sequence 7, Appl1
42	385.5	24.1	1421	US-10-014-717-2	Sequence 2, Appl1
43	346	21.7	328	US-09-735-056-31	Sequence 31, Appl1
44	317.5	19.9	2563	US-09-836-705-46	Sequence 46, Appl1
45	317	19.8	1610	US-09-738-626-6666	Sequence 6666, Ap

ALIGNMENTS

RESULT 1
US-09-808-880-3
Sequence 3, Application US/09808880
Publication No. US20030027287A1
GENERAL INFORMATION:
APPLICANT: Beclach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/808,880
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: US/09/428,517
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/120,254
PRIOR FILING DATE: 1999-02-16
PRIOR APPLICATION NUMBER: 60/106,100
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ IDS NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 3816
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: Oleandolide PKS
US-09-808-880-3

Query Match 55.8%; Score 891.5; DB 9; Length 3816;
Best Local Similarity 58.6%; Pred. No. 2.1e-71;
Matches 180; Conservative 31; Mismatches 93; Indels 3; Gaps 2;
QY 1 VFVFPQGGQGMAGVRLAASSPVFARRMQACEBALAPWVDSVVDLRRDAGDAVERA 60
DB 567 VFVFPQGGQGMAGVRLAASSPVFARRMQACEBALAPWVDSVVDLRRDAGDAVERA 60
QY 61 DVVQPLVFTSMVSLALMSYGLIEPDVAVLGHSGCEIAAAVCAALSLKDAKTVAAVBSRA 120
DB 625 DVVQPLVFTSMVSLALMSYGLIEPDVAVLGHSGCEIAAAVCAALSLKDAKTVAAVBSRA 120

QY	121	LAARVRGGMAVSPLPAQEVTEOLIGRRMAGRLMVAANAGPSTAAVSGPAEAVDEVLAYCA	160	
		685	IAIIAGGGMVASVSLPAGRVRTVL-DTYGRLSYAAVNGPSTVVSQDAQMLDELGCE	743
Db	181	GTGVRAARRIPVDYASHCPHVQPLREBLLLELDGISPOESGVFFSTVEGTLDDPTTLDDAA	240	
QY	744	REGVRAARVVDVAHSHSAQMDQLRDELLEALADITPQSSVVFSTVADWLDTTALDAG	803	
Db	241	YWRRLNHQPRFSDAYQALADDGHRVYVAVSHPLTVLAIEDTTEEDTAEVDYTAIGSLRRG	300	
QY	804	YMFNTLRRETVRFOEAVEAGLVQAQMGAFVFCSPHPLVTVGIEQTLDITVEADAVALLGSLRRD	863	
Db	301	DNDTRRF	307	
QY		:		
Db	864	EGGLGRF	870	

```

RESULT 2
US-09-808-880-2
: Sequence 2, Application US/09808880
: Publication No. US20030027287A1
: GENERAL INFORMATION:
: APPLICANT: Betlach, Mary C.
: APPLICANT: Shah, Sanjay Krishnakant
: APPLICANT: McDaniel, Robert
: APPLICANT: Tang, Li
: TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
: FILE REFERENCE: 30062-20029.00
: CURRENT APPLICATION NUMBER: US/09/808,880
: CURRENT FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: US/09/428,517
: PRIOR FILING DATE: 1998-10-28
: PRIOR APPLICATION NUMBER: 60/120,254
: PRIOR FILING DATE: 1999-02-16
: PRIOR APPLICATION NUMBER: 60/106,100
: PRIOR FILING DATE: 1998-10-29
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: LENGTH: 4150
: TYPE: PR1
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Recombinant
: OTHER INFORMATION: Oleandolide PRS
: US-09-808-880-2

```

Query Match	53.9%;	Score 861.5;	DB 9;	Length 4150;
Best Local Similarity	56.7%;	Pred. No. 1.2e-68;		
Matches 174;	Conservative. .35;	Mismatches 95;	Indels 3;	Gaps 2;

QY	1	VFPVPGGCGXGAAGKGVRLTLASSPFARBMQACEELAAWVMVSVDYDILRRAGDGVWERA	60
Db	3076	VFPVPGGCGXGVGAAGGLDLACPFPAEVAECAATLDELTKGSLVEYR--GGEAVLGRV	3133
QY	61	DVVPVPLFSVNVSLAALMRSYGIEPDAVLVHSGGEIATAAHVCGALSLDKAAKTVALARSA	120
Db	3134	DVVPALMAVWVSLARTWRYYGVPEAAVAVHSGGEIATAACVAGGISLADGARVVTLARSA	3192
QY	121	LAARGRGMAVSPLEPAQVEEQILGERVAGRLMTVAANVGPRSTAVSGDAEAVDEVLAACA	180
Db	3194	IARLAGGGMVSVGLSARVATML-DYTGKRVSAVANVGPSSTVSSGQAQLDELLACGE	3252
QY	181	GTGVPARRRIPVDVASHCHVOPRLREELIELLDGISPPQSGVPPFSTVEGTWLDITTLDA	240
Db	3253	REGVARRARVPADVASHSAQMQLRDELLEALADVPPQDSSVFFSTVADMDITTLADAG	3312
QY	241	YMYNLTLPVRFSAVQALADGGHVFVFSVPHPTLVPAIEDTDTDAEDVTALGSLRG	300
Db	3313	YMFNLTARTVAFQEAIVEGLVAQGMCAFECCSPHPLVYGIETLDTLPADANVALSLSRD	3372
QY	301	DNDTRRF	307

Db 3373 EGGLDRF 3379

```

RESULT 3
US-09-808-880-4
Sequence 4, Application US/09808880
Publication No. US20030027287a1
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029..00
CURRENT APPLICATION NUMBER: US/09/808, 880
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: US/09/428, 517
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/120,254
PRIOR FILING DATE: 1999-02-16
PRIOR APPLICATION NUMBER: 60/106,100
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 3519
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-808-880-4

```

Query Match 858; DB 9; Length 3519;
 Best Local Similarity 57.0%; Pred. No. 25:68;
 Matches 175; Conservative 34; Mismatches 94; Indels 4; Gaps 3;

QY	1	VFPVPPGGGGMMAGGVRVTLASSPFEARRMQACEALAPWDMKSVVDLIR--RDADGDAWVER	59
Db	2221	VFPVPPGGGGMMVGMAAGLLDACPFAPAAVABECAAALDPVTGMSLVEVYLQGHDA--TVLGR	2278
QY	60	ADVVPQPLFSVMVSLAALMRSYGIEPDAVLGHSGEIAAAHVCGALSILKDAKTVALRSR	119
Db	2279	VDVVPRLMAVMNLSLARTWRYGYVEPAVAVGHSGEILAAACVAGGSLADQARVVLRSR	2338
QY	120	ALAAVRRGGGMASVPLPAQVEQLIGERMAGRLVWAAVNGPRSTAVSGDAEAVDEVILAYC	179
Db	2339	ALIRAGGGGMVSVSLPAGRVRTML-DTYGGRVSVAAVNGPSSITVSGDVQALDELLAGC	2397
QY	180	AGTGVTRARRIPVDYASHCPHVQPLREELBELLDLSHQSGVFPFFSTVEGTLDDTLTIDA	239
Db	2398	EREGRVARRRPVVDYASHSAQMQLRDELLEALADITPODSSVFPFFSTVADMDLTTMLDA	2457
QY	240	AYMYRNLMHQVREPSDAVQALADDGHRVFEVSPHPTLVPAILEDTEETAEDVTAIGSLRR	299
Db	2458	GYMFTNLRVTRVQEAIVEGLVAQGMGAFEVCSPPVLVPGIEQLDLDADNAVALGSLRR	2517
QY	300	GDNDTRR	306
Db	2518	DEGGDLR	2524

US-RESULT 4
 09-73735-056-33
 : Sequence 33, Application US/09735056
 : Publication No. US20090013662A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Katz, Leonard
 : APPLICANT: Stasasi, Diane L.
 : APPLICANT: Summers Jr., Richard G
 : APPLICANT: Ruan, Xiaolan
 : APPLICANT: Pereda-Lopez, Ana
 : APPLICANT: Kakavava, Stephan J.

Db 562 FVPPGGGTQWAGMGBELDSSPEFAAMACECTALSPYDWSLEAVVRQAPSAPTLDRVD 621

QY 62 VVQVPLFSVWWSLAALMRSYGIPEPDAVLGHSQGEIAAAHVCGALSLDKAAKTVALRSRAL 121

Db 622 VVQVPLFAVWWSLAKVQHGHGITPEAVIGHSQGEIAAAVYAGALTLDAAKRVTLRSKSI 681

QY 122 AA-VRGRGMAVPL-----PAQEVETOLIGERWAGRLVVAANGPRSTAVSGDAEAVDEVL 176

Db 682 AAHLAGKGMISLASEBATRQRIENLHG-----LSIAAVNGPTATVSGDPTQIQELA 735

QY 177 AYCAGTGVARRIIPVDYASHCHPVQPLREBELLELDISQPSGVPPFSTVEGTWLDTTT 236

Db 736 QACADGIRARIIPVDYASHASHAVETIENELADVLAGISPTQPVPPFSTLEGWTITEPA 795

QY 237 LDAAYWYRNLIHQVRFSDAVQALA-DDGHRVFEVSPHPTLVPAIETTEDTAEDVTAIG 295

Db 796 LDGGYWRNLIHRVGFAPAVETLATDEGTFHFIEVSAHPVLTMTLPD-----KVTGLA 848

QY 296 SLRRGNDTTR 306

Db 849 TLRRDGGQHR 859

RESULT 7

US-09-836-821-37

Sequence 37, Application US/09836821

Publication No. US20030087405A1

GENERAL INFORMATION:

APPLICANT: Sherman, D.H.

APPLICANT: Liu, H.

APPLICANT: Xue, Y.

APPLICANT: Zhao, L.

TITLE OF INVENTION: DNA encoding methymycin and pikromycin

FILE REFERENCE: 600.438US1

CURRENT APPLICATION NUMBER: US/09/836,821

PRIOR FILING DATE: 2001-04-17

PRIOR APPLICATION NUMBER: 09/105,537

PRIOR FILING DATE: 1998-06-26

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 37

LENGTH: 1346

TYPE: PRT

ORGANISM: Streptomyces venezuelae

US-09-836-821-37

Query Match 48.3%; Score 772.5; DB 9; Length 1346;

Best Local Similarity 51.4%; Pred. No. 3.2e-61;

Matches 160; Conservative 42; Mismatches 90; Indels 19; Gaps 5;

QY 2 FVPPGGGTQWAGMGBELDSSPEFAAMACECTALSPYDWSLEAVVRQAPSAPTLDRVD 61

Db 562 FVPPGGGTQWAGMGBELDSSPEFAAMACECTALSPYDWSLEAVVRQAPSAPTLDRVD 621

QY 62 VVQVPLFSVWWSLAALMRSYGIPEPDAVLGHSQGEIAAAHVCGALSLDKAAKTVALRSRAL 121

Db 622 VVQVPLFAVWWSLAKVQHGHGITPEAVIGHSQGEIAAAVYAGALTLDAAKRVTLRSKSI 681

QY 122 AA-VRGRGMAVPL-----PAQEVETOLIGERWAGRLVVAANGPRSTAVSGDAEAVDEVL 176

Db 682 AAHLAGKGMISLASEBATRQRIENLHG-----LSIAAVNGPTATVSGDPTQIQELA 735

QY 177 AYCAGTGVARRIIPVDYASHCHPVQPLREBELLELDISQPSGVPPFSTVEGTWLDTTT 236

Db 736 QACADGIRARIIPVDYASHASHAVETIENELADVLAGISPTQPVPPFSTLEGWTITEPA 795

QY 237 LDAAYWYRNLIHQVRFSDAVQALA-DDGHRVFEVSPHPTLVPAIETTEDTAEDVTAIG 295

Db 796 LDGGYWRNLIHRVGFAPAVETLATDEGTFHFIEVSAHPVLTMTLPD-----KVTGLA 848

QY 296 SLRRGNDTTR 306

Db 849 TLRRDGGQHR 859

RESULT 8

US-09-793-708-4

Sequence 4, Application US/09793708

Publication No. US20030104597A1

GENERAL INFORMATION:

APPLICANT: ASHLEY, Gary

APPLICANT: BETLACH, Melanie C.

APPLICANT: BETLACH, Mary C.

APPLICANT: MCDANIEL, Robert

APPLICANT: TANG, Li

TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE

FILE REFERENCE: 300622002121

CURRENT APPLICATION NUMBER: US/09/793,708

PRIOR FILING DATE: 2001-02-22

PRIOR APPLICATION NUMBER: US 09/657,440

PRIOR FILING DATE: 2000-09-07

PRIOR APPLICATION NUMBER: US 09/320,878

PRIOR FILING DATE: 1999-05-27

PRIOR APPLICATION NUMBER: US 09/141,908

PRIOR FILING DATE: 1998-08-28

PRIOR APPLICATION NUMBER: US 09/073,538

PRIOR FILING DATE: 1998-05-06

PRIOR APPLICATION NUMBER: US 08/846,247

PRIOR FILING DATE: 1997-04-30

PRIOR APPLICATION NUMBER: US 60/134,990

PRIOR FILING DATE: 1999-05-20

NUMBER OF SEQ ID NOS: 38

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 4

LENGTH: 1346

TYPE: PRT

ORGANISM: Streptomyces venezuelae

US-09-793-708-4

Query Match 48.3%; Score 772.5; DB 9; Length 1346;

Best Local Similarity 51.4%; Pred. No. 3.2e-61;

Matches 160; Conservative 42; Mismatches 90; Indels 19; Gaps 5;

QY 2 FVPPGGGTQWAGMGBELDSSPEFAAMACECTALSPYDWSLEAVVRQAPSAPTLDRVD 61

Db 562 FVPPGGGTQWAGMGBELDSSPEFAAMACECTALSPYDWSLEAVVRQAPSAPTLDRVD 621

QY 62 VVQVPLFSVWWSLAALMRSYGIPEPDAVLGHSQGEIAAAHVCGALSLDKAAKTVALRSRAL 121

Db 622 VVQVPLFAVWWSLAKVQHGHGITPEAVIGHSQGEIAAAVYAGALTLDAAKRVTLRSKSI 681

QY 122 AA-VRGRGMAVPL-----PAQEVETOLIGERWAGRLVVAANGPRSTAVSGDAEAVDEVL 176

Db 682 AAHLAGKGMISLASEBATRQRIENLHG-----LSIAAVNGPTATVSGDPTQIQELA 735

QY 177 AYCAGTGVARRIIPVDYASHCHPVQPLREBELLELDISQPSGVPPFSTVEGTWLDTTT 236

Db 736 QACADGIRARIIPVDYASHASHAVETIENELADVLAGISPTQPVPPFSTLEGWTITEPA 795

QY 237 LDAAYWYRNLIHQVRFSDAVQALA-DDGHRVFEVSPHPTLVPAIETTEDTAEDVTAIG 295

Db 796 LDGGYWRNLIHRVGFAPAVETLATDEGTFHFIEVSAHPVLTMTLPD-----KVTGLA 848

QY 296 SLRRGNDTTR 306

Db 849 TLRRDGGQHR 859

RESULT 9

US-09-861-289-37

Sequence 37, Application US/09861289

Patent No. US20020110897A1

GENERAL INFORMATION:

APPLICANT: Sherman, D.H.

APPLICANT: Liu, H.

APPLICANT: Xue, Y.

APPLICANT: Zhao, L.
FILE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438US1
CURRENT APPLICATION NUMBER: US/09/861,289
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 37
LENGTH: 1346
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-861-289-37

Query Match 48.3%; Score 772.5; DB 10; Length 1346;
Best Local Similarity 51.4%; Pred. No. 3.2e-61;
Matches 160; Conservative 42; Mismatches 90; Indels 19; Gaps 5;

QY 2 FVFPGGCGWAGNGVRLASGPFVARRMOACEALAPWVDSVVDILRRDAGDAWVERAD 61
DB 562 FVFPGGCGTOWAGNGABLLDSSPEFAAAMACETALSTPYVDMSLAVVRQAAPATLDRVD 621
QY 62 VVQPVLFVWVSLAALMRSYGIPEPDVVLGHSOGEIAAAHYCGALSLKDAKTVALLRSRAL 121
DB 622 VVQPVTFVWVSLAKWQHGHGVTPEAVYIGHSOGEIAAAYAGALTLDAAARVVTLRKSKI 681
QY 122 AA-VRGRGMAVPLPAQVEQLIGERWAG--RLWVAVNGPSTAVSGDAEAVDEVL 176
DB 682 AAHLAAGGGMISLASEATRORLENLHG-----LSIAANGTATVVGSDPTQIOELAA 735
QY 177 AYCAGTVARRIIPVDYASHCHPVPRLBELLELGDISPQSGVPFSTVEGTWLDTTT 236
DB 736 QACGADGIRARIIPVDYASHAHVETIENELADVLGSLSPQTPVPFSTLEGWIREPA 795
QY 237 LDAAYRNLRHOPVRFSDAVQALA-DGHRVFEVSPHPTLVPAIEDTTEDTAEDVTAIG 295
DB 796 LDGQWYRNLRHVRGFAVAVETLATDEGFTHFVSAHPVLTMLPD-----KVTGLA 848
QY 296 SLRGRDNDTRR 306
DB 849 TLRREDGGQHR 859

RESULT 10
US-09-860-846-33

Sequence 33, Application US/09860846
Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438US1
CURRENT APPLICATION NUMBER: US/09/860,846
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-860-846-33

Query Match 48.3%; Score 772.5; DB 9; Length 3739;
Best Local Similarity 53.1%; Pred. No. 1.2e-60;
Matches 164; Conservative 38; Mismatches 92; Indels 15; Gaps 5;
QY 2 FVFPGGCGWAGNGVRLASGPFVARRMOACEALAPWVDSVVDILRRDAGDAWVERAD 61

DB 2069 FVFPGGCGTOWAGNGABLLDVSKEFAAAMECEALAPYVDWSLEAVVRQAPGPTLERVD 2128
QY 62 VVQPVLFVWVSLAALMRSYGIPEPDVVLGHSOGEIAAAHYCGALSLKDAKTVALLRSRAL 121
DB 2129 VVQPVTFVWVSLAKWQHGHGVTPEAVYIGHSOGEIAAAYAGALTLDAAARVVTLRKSKI 2188
QY 122 AA-VRGRGMAVPLPAQVEQLIGERWAG--RLWVAVNGPSTAVSGDAEAVDEVLAY 178
DB 2189 GAHLAAGGGMISLASEATRORLENLHG-----LSIAANGTATVVGSDPTQIOELAA 2244
QY 179 CAGTVARRIIPVDYASHCHPVPRLBELLELGDISPQSGVPFSTVEGTWLDTTLD 238
DB 2245 CEADGVARRIIPVDYASHAHVETIENELADVLGSLSPQTPVPFSTLEGWIREPALD 2304
QY 239 AAYWYRNLRHOPVRFSDAVQALA-DGHRVFEVSPHPTLVPAIEDTTEDTAEDVTAIGSL 297
DB 2305 GGTYWYRNLRHVRGFAVAVETLATDEGFTHFVSAHPVLTMLPET-----VTGLGTL 2357
QY 298 RRGDNDTRR 306
DB 2358 RRDNGGQHR 2366

RESULT 11
US-09-988-384B-33
Sequence 33, Application US/0988384B
Publication No. US2003007382A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,536US1
CURRENT APPLICATION NUMBER: US/09/988,384B
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: PCT/US99/14398
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 53
SEQ ID NO 33
LENGTH: 3739
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-988-384B-33

Query Match 48.3%; Score 772.5; DB 9; Length 3739;
Best Local Similarity 53.1%; Pred. No. 1.2e-60;
Matches 164; Conservative 38; Mismatches 92; Indels 15; Gaps 5;

QY 2 FVFPGGCGWAGNGVRLASGPFVARRMOACEALAPWVDSVVDILRRDAGDAWVERAD 61
DB 2069 FVFPGGCGTOWAGNGABLLDVSKEFAAAMECEALAPYVDWSLEAVVRQAPGPTLERVD 2128
QY 62 VVQPVLFVWVSLAALMRSYGIPEPDVVLGHSOGEIAAAHYCGALSLKDAKTVALLRSRAL 121
DB 2129 VVQPVTFVWVSLAKWQHGHGVTPEAVYIGHSOGEIAAAYAGALTLDAAARVVTLRKSKI 2188
QY 122 AA-VRGRGMAVPLPAQVEQLIGERWAG--RLWVAVNGPSTAVSGDAEAVDEVLAY 178
DB 2189 GAHLAAGGGMISLASEATRORLENLHG-----LSIAANGTATVVGSDPTQIOELAA 2244
QY 179 CAGTVARRIIPVDYASHCHPVPRLBELLELGDISPQSGVPFSTVEGTWLDTTLD 238
DB 2245 CEADGVARRIIPVDYASHAHVETIENELADVLGSLSPQTPVPFSTLEGWIREPALD 2304
QY 239 AAYWYRNLRHOPVRFSDAVQALA-DGHRVFEVSPHPTLVPAIEDTTEDTAEDVTAIGSL 297
DB 2305 GGTYWYRNLRHVRGFAVAVETLATDEGFTHFVSAHPVLTMLPET-----VTGLGTL 2357
QY 298 RRGDNDTRR 306

Db 2358 RRDNGGQHR 2366

RESULT 12

US-09-836-821-33
; Sequence 33, Application US/09836821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 3739
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-836-821-33

Query Match 48.3%; Score 772.5; DB 9; Length 3739;

Best Local Similarity 53.1%; Pred. No. 1.2e-60;

Matches 164; Conservative 38; Mismatches 92; Indels 15; Gaps 5;

QY 2 FVFPGGGQWAGVRLASSPVFARRMQACEALAPWDMVVDILRRDAGDAVWERAD 61

Db 2069 FVFPGGGQWAGVRLASSPVFARRMQACEALAPWDMVVDILRRDAGDAVWERAD 2128

QY 62 VVQPVLFVSWSLAALWRSYGIPEPDVLTGHSQGEIAAAHVCGALSLDKAAVTALSRAL 121

Db 2129 VVQPVTFVAVVWSLAKWQHGHGVTPOAVVGHSGGEIAAAVYAGALSLDAAVVTLSKSI 2188

QY 122 AA-VRGGGMAVPLPAQVEQLIGERWAG--RLWVAAVNGPRSTAVSGDAEAVDEVLAY 178

Db 2189 GAHLAAGGGMVLSLALS---EAAVVERLAGFDGLSVAAVNGPRATVVSDDPTQIOELAOA 2244

QY 179 CAGTGVARRIPVDYASHCPVQPLREBELLELDISPOSGVPFFSTVEGTWLDITTTLD 238

Db 2245 CEADGVARRIPVDYASHSHAVETISELADVLAGLSPTQVFFSTLBCAWITTEPALD 2304

QY 239 AAYWYRNLHQPVRPSDAVQALA-DDGHRVFEVSPHPTLVPAIDTTEDTAEDVTAIGSL 297

Db 2305 GGWYWRNLRRHVGEPAPVETLATDEGFTHFVEVSAHPVLTWALPET-----VTGLGTL 2357

QY 298 RRGDNDTRR 306

Db 2358 RRDNGGQHR 2366

RESULT 13

US-09-793-708-2

; Sequence 2, Application US/09793708

; Publication No. US20030104597A1

; GENERAL INFORMATION:

; APPLICANT: Ashley, Gary

; APPLICANT: BETLACH, Melanie C.

; APPLICANT: BETLACH, Mary C.

; APPLICANT: MCDANIEL, Robert

; APPLICANT: TANG, Li

; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE

; FILE REFERENCE: 300622002121

; CURRENT APPLICATION NUMBER: US/09/793,708

; CURRENT FILING DATE: 2001-02-22

; PRIOR APPLICATION NUMBER: US 09/657,440

; PRIOR FILING DATE: 2000-09-07

; PRIOR APPLICATION NUMBER: US 09/320,878

; PRIOR FILING DATE: 1999-05-27

; PRIOR APPLICATION NUMBER: US 09/141,908

; PRIOR FILING DATE: 1998-08-28

; PRIOR APPLICATION NUMBER: US 09/073,538

; PRIOR FILING DATE: 1998-05-06

; PRIOR APPLICATION NUMBER: US 08/846,247

; PRIOR FILING DATE: 1997-04-30

; PRIOR APPLICATION NUMBER: US 60/134,990

; PRIOR FILING DATE: 1999-05-20

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 3739

; TYPE: PRT

; ORGANISM: Streptomyces venezuelae

US-09-793-708-2

Query Match 48.3%; Score 772.5; DB 9; Length 3739;

Best Local Similarity 53.1%; Pred. No. 1.2e-60;

Matches 164; Conservative 38; Mismatches 92; Indels 15; Gaps 5;

QY 2 FVFPGGGQWAGVRLASSPVFARRMQACEALAPWDMVVDILRRDAGDAVWERAD 61

Db 2069 FVFPGGGQWAGVRLASSPVFARRMQACEALAPWDMVVDILRRDAGDAVWERAD 2128

QY 62 VVQPVLFVSWSLAALWRSYGIPEPDVLTGHSQGEIAAAHVCGALSLDKAAVTALSRAL 121

Db 2129 VVQPVTFVAVVWSLAKWQHGHGVTPOAVVGHSGGEIAAAVYAGALSLDAAVVTLSKSI 2188

QY 122 AA-VRGGGMAVPLPAQVEQLIGERWAG--RLWVAAVNGPRSTAVSGDAEAVDEVLAY 178

Db 2189 GAHLAAGGGMVLSLALS---EAAVVERLAGFDGLSVAAVNGPRATVVSDDPTQIOELAOA 2244

QY 179 CAGTGVARRIPVDYASHCPVQPLREBELLELDISPOSGVPFFSTVEGTWLDITTTLD 238

Db 2245 CEADGVARRIPVDYASHSHAVETISELADVLAGLSPTQVFFSTLBCAWITTEPALD 2304

QY 239 AAYWYRNLHQPVRPSDAVQALA-DDGHRVFEVSPHPTLVPAIDTTEDTAEDVTAIGSL 297

Db 2305 GGWYWRNLRRHVGEPAPVETLATDEGFTHFVEVSAHPVLTWALPET-----VTGLGTL 2357

QY 298 RRGDNDTRR 306

Db 2358 RRDNGGQHR 2366

RESULT 14

US-09-861-289-33

; Sequence 33, Application US/09861289

; Patent No. US20020110897A1

; GENERAL INFORMATION:

; APPLICANT: Sherman, D.H.

; APPLICANT: Liu, H.

; APPLICANT: Xue, Y.

; APPLICANT: Zhao, L.

; TITLE OF INVENTION: DNA encoding methymycin and pikromycin

; FILE REFERENCE: 600.438US1

; CURRENT APPLICATION NUMBER: US/09/861,289

; CURRENT FILING DATE: 2001-05-18

; PRIOR APPLICATION NUMBER: 09/105,537

; PRIOR FILING DATE: 1998-06-26

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 33

; LENGTH: 3739

; TYPE: PRT

; ORGANISM: Streptomyces venezuelae

US-09-861-289-33

Query Match 48.3%; Score 772.5; DB 10; Length 3739;

Best Local Similarity 53.1%; Pred. No. 1.2e-60;

Matches 164; Conservative 38; Mismatches 92; Indels 15; Gaps 5;

QY 2 FVFPGGGQWAGVRLASSPVFARRMQACEALAPWDMVVDILRRDAGDAVWERAD 61

Db 7307 RRDNGGQHR 7315
Search completed: June 17, 2003, 13:46:41
Job time : 10.5655 secs

Db 2069 FVPPGQCTQWAGCAELLDVSKFPAAMACEALAPYVMSLEAVVROAPGAPTLERVD 2128
QY 62 VVQPVLFVSVVSLAALRSYGIEPDAVLGHSQGEIAAAYGALSLDKAAKTVALRSRAL 121
Db 2129 VVQPVTFVAVVSLAKVQHGHGVTPOAVGHSQGEIAAAYGALSLDDAARVVTLRKSKI 2188
QY 122 AA-VRGRGNASVPLPAQVEEQLIGERMAG--RLMVAAVNGPRSTAVSGDAEAVDEVLAY 178
Db 2189 GAHLAGGGGMLSLALS---EAAVERLAGFDGLSVAAVNGPTATVSGDPTQIOELAOA 2244
QY 179 CAGTGVARARIIPVDYASHCHVOPRLRELLLELDISPPSGVPPFSTVEGTWLDTTLD 238
Db 2245 CEADGVARARIIPVDYASHSAHVERISELADVLGSLPQTPQVFPFSTLEGAMITEPALD 2304
QY 239 AAYWYRLHQVPVPSDAVQALA--DDGHRVFEVSPHPTLVPAIEDTETAEDVTAIGSL 297
Db 2305 GGTYWYRLRHRVGPAPAVETLATDEGPTHFEVSAHPVLTMALPET-----VTGLGTL 2357
QY 298 RRGDNDTRR 306
Db 2358 RRDNGGQHR 2366

RESULT 15
US-09-860-846-6
; Sequence 6, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 11877
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-6

Query Match 48.3%; Score 772.5; DB 9; Length 11877;
Best Local Similarity 53.1%; Pred. No. 5,7e-60;
Matches 164; Conservative 38; Mismatches 92; Indels 15; Gaps 5;

QY 2 FVPPGQCTQWAGCAELLDVSKFPAAMACEALAPYVMSLEAVVROAPGAPTLERVD 61
Db 7018 FVPPGQCTQWAGCAELLDVSKFPAAMACEALAPYVMSLEAVVROAPGAPTLERVD 7077
QY 62 VVQPVLFVSVVSLAALRSYGIEPDAVLGHSQGEIAAAYGALSLDKAAKTVALRSRAL 121
Db 7078 VVQPVTFVAVVSLAKVQHGHGVTPOAVGHSQGEIAAAYGALSLDDAARVVTLRKSKI 7137
QY 122 AA-VRGRGNASVPLPAQVEEQLIGERMAG--RLMVAAVNGPRSTAVSGDAEAVDEVLAY 178
Db 7138 GAHLAGGGGMLSLALS---EAAVERLAGFDGLSVAAVNGPTATVSGDPTQIOELAOA 7193
QY 179 CAGTGVARARIIPVDYASHCHVOPRLRELLLELDISPPSGVPPFSTVEGTWLDTTLD 238
Db 7194 CEADGVARARIIPVDYASHSAHVERISELADVLGSLPQTPQVFPFSTLEGAMITEPALD 7253
QY 239 AAYWYRLHQVPVPSDAVQALA--DDGHRVFEVSPHPTLVPAIEDTETAEDVTAIGSL 297
Db 7254 GGTYWYRLRHRVGPAPAVETLATDEGPTHFEVSAHPVLTMALPET-----VTGLGTL 7306
QY 298 RRGDNDTRR 306

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 12:54:32 ; Search time 5.57989 Seconds
(without alignments)
1618,821 Million cell updates/sec

Title: US-09-914-286-3_COPY_1050_1356
Perfect score: 1598
Sequence: 1 VFVFPQGGQGMAGVRLA.....AEDVTAIGLRGNDTRRF 307

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	891.5	55.8	3816	US-09-428-517-3	Sequence 3, Appl
2	884.5	55.4	3567	US-07-642-734C-4	Sequence 4, Appl
3	884.5	55.4	3567	US-08-439-009A-4	Sequence 4, Appl
4	861.5	53.9	4150	US-09-428-517-2	Sequence 2, Appl
5	858	53.7	3519	US-09-428-517-4	Sequence 2, Appl
6	852	53.3	2595	US-08-036-987A-2	Sequence 2, Appl
7	852	53.3	2595	US-09-370-700-2	Sequence 2, Appl
8	848.5	53.1	3170	US-09-036-987A-4	Sequence 2, Appl
9	848.5	53.1	3170	US-09-370-700-4	Sequence 2, Appl
10	840.5	52.6	3491	US-07-642-734C-2	Sequence 4, Appl
11	840.5	52.6	3491	US-08-439-009A-2	Sequence 2, Appl
12	828	51.8	344	US-08-858-003-33	Sequence 3, Appl
13	828	51.8	344	US-09-078-166-33	Sequence 3, Appl
14	828	51.8	344	US-08-997-467-33	Sequence 3, Appl
15	822.5	51.5	3724	US-08-804-227C-10	Sequence 10, Appl
16	822.5	51.5	3724	US-08-804-198-4	Sequence 10, Appl
17	816.5	51.1	829	US-09-413-814-105	Sequence 105, App
18	814.5	51.0	3170	US-07-642-734C-5	Sequence 5, Appl
19	814.5	51.0	3170	US-08-439-009A-5	Sequence 5, Appl
20	774	48.4	5588	US-09-036-987A-6	Sequence 6, Appl
21	774	48.4	5588	US-09-370-700-6	Sequence 6, Appl
22	773	48.4	1580	US-08-804-227C-11	Sequence 11, Appl
23	773	48.4	1580	US-08-804-198-5	Sequence 11, Appl
24	772.5	48.3	1346	US-09-330-878-4	Sequence 4, Appl
25	772.5	48.3	1346	US-09-105-537-37	Sequence 37, Appl
26	772.5	48.3	3739	US-09-320-878-2	Sequence 2, Appl
27	772.5	48.3	3739	US-09-105-537-33	Sequence 33, Appl

28	772.5	48.3	11877	US-09-105-537-6	Sequence 6, Appl
29	772	48.3	345	US-08-858-003-34	Sequence 34, Appl
30	772	48.3	345	US-09-078-166-34	Sequence 34, Appl
31	772	48.3	345	US-08-997-467-34	Sequence 34, Appl
32	769	48.1	4551	US-09-320-878-1	Sequence 1, Appl
33	769	48.1	4613	US-09-105-537-31	Sequence 31, Appl
34	763	47.7	1562	US-09-320-878-3	Sequence 3, Appl
35	763	47.7	1562	US-09-105-537-35	Sequence 35, Appl
36	754	47.2	3729	US-08-804-227C-4	Sequence 4, Appl
37	754	47.2	3729	US-08-804-227C-2	Sequence 2, Appl
38	733	45.9	1611	US-08-804-227C-5	Sequence 5, Appl
39	733	45.9	4545	US-08-804-227C-14	Sequence 14, Appl
40	732.5	45.8	7257	US-09-335-409-5	Sequence 5, Appl
41	732.5	45.8	7257	US-09-568-102-5	Sequence 5, Appl
42	732.5	45.8	7257	US-09-567-969-5	Sequence 5, Appl
43	732.5	45.8	7257	US-09-568-480-5	Sequence 5, Appl
44	732.5	45.8	7257	US-09-568-486-5	Sequence 5, Appl
45	732.5	45.8	7257	US-09-568-472-5	Sequence 5, Appl

ALIGNMENTS

```

RESULT 1
US-09-428-517-3
; Sequence 3, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3816
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
; OTHER INFORMATION: Oleandolide PKS
US-09-428-517-3
Query Match      55.8%; Score 891.5; DB 4; Length 3816;
Best Local Similarity 58.6%; Pred. No. 5.7e-81;
Matches 180; Conservative 31; Mismatches 93; Indels 3; Gaps 2;

QY 1 VFVFPQGGQGMAGVRLAASPVPARBMQACEBALAPWDMGVSDTLRRDAGAVVERA 60
DB 567 VFVFPQGGQGMAGVRLAACPVPFAVAECAVAIDPLTGMVLVEYL--GGEAVLGRV 624
QY 61 DVVQPVLFVYVSLAALMRSYGIPEPAVLGSGGEIAAAHVCAALSLKDAKTVLARSRA 120
DB 625 DVVQPVLMAMVSLATWRYGVEPAVAVGSGGEIAAACVAGGLSLADGARVVLARSRA 684
QY 121 LAAVRGCGMAASP/LPAQOEYEOJLIGERMAGRLVAAVNGRSTAVSGDAVAVBYLAC 180
DB 685 IARIAGCGGVSVSLPAGRVRTWL-DYGGRLSVAAVNGPSSVTVSGDAQALDELACGE 743
QY 181 GTGVARRRIPVDVASHCPHVPQRLREELLETGLDISPOSGVPFSTVEGTWLTDTTLDAA 240
DB 744 REGVRRRVVDVASHAGMDQLRDELLEALADITQHSSVPPFSTVTADMTTALDAG 803
QY 241 YWRNLHQPVRFSDAVQALADGHRVFEVSPPTLVPAIEDTETDAEDVTAIGSLRNG 300

```

Db 804 YWFTNLRETVRFOEAVEGLVAQGMGAFVCSPPHVLVPGIEQTLDTVEADAVALSRRD 863
QY 301 DNDTRRF 307
Db 864 EGGIGRF 870

RESULT 2

US-07-642-734C-4
; Sequence 4, Application US/07642734C
; Patent No. 5824513
; GENERAL INFORMATION:
; APPLICANT: Katz, L
; APPLICANT: Donadio, S
; APPLICANT: Mcalpine, J B
; TITLE OF INVENTION: Recombinant DNA Method for Producing
; Erythromycin Analogs
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward H. Gorman
; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
; CITY: Abbott Park
; STATE: IL
; COUNTRY: US
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/642,734C
; FILING DATE: 17-JAN-91
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckerts, Andreas M
; REGISTRATION NUMBER: 32652
; REFERENCE/DOCKET NUMBER: 4952, US. 01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-9396
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3567 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-642-734C-4

Query Match 55.4%; Score 884.5; DB 2; Length 3567;
Best Local Similarity 57.3%; Pred. No. 2.6e-80;
Matches 177; Conservative 41; Mismatches 88; Indels 3; Gaps 2;
QY 1 VFVPPGGGGMAGNGVRLASSPVFARRMQACEBALAPWVDSVVDILRRDAGDAVERA 60
Db 560 VFLEPGGSGMAGNGAELLSSPVFAGKIRACDSMA PMQDWKXSVLRRQAPAGLDIV 619
QY 61 DVVQPVLFSTWVSLAALMRSYGIIPDAVLGHSGCEIAAAHVCGALSLKDAKTVALSRA 120
Db 620 DVVQPVLFSTWVSLAALMRSYGIIPDAVLGHSGCEIAAAHVCGALSLKDAKTVALSRA 120
QY 121 LAAVRGGGMAVPLPAOEVEQLIGERWAGRLWVAANGPRSTAVSGDAEAVDEVLA YCA 180
Db 680 MRSLSGGGMAVPLPAOEVEQLIGERWAGRLWVAANGPRSTAVSGDAEAVDEVLA YCA 180
QY 181 GTGVARRIRPVYASHCNVPOLREBELLELIGDISPPSGVFPFSTYEGTWLDTTITDAA 240
Db 739 AEGIRVADIDVDYASHSPQIERVREBELLETTGDIAPRPARVTFHSTVESRMDGTELDAR 798
QY 241 YWYRNHLQPVRFSDAVOALADGHRVFEVSPHPTLVPAIEDTTE--DTAEDVTAIGSLR 298
Db 799 YWYRNHLRETVRFADAVTRLAESGYDAFIEVSPHVVVQAVEAEVADGAEADAVVVGSLH 858

QY 299 RGNDRTRF 307
Db 859 RDGGDLASF 867

RESULT 3

US-08-439-009A-4
; Sequence 4, Application US/08439009A
; Patent No. 6004787
; GENERAL INFORMATION:
; APPLICANT: Donadio, S
; APPLICANT: Katz, L
; APPLICANT: Mcalpine, J B
; TITLE OF INVENTION: Method of Directing Biosynthesis of
; Specific Polypeptides
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven F. Weinstein
; STREET: Abbott Laboratories D377/AP6D-2 One Abbott
; CITY: Abbott Park
; STATE: IL
; COUNTRY: US
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,009A
; FILING DATE: 11-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Casuto, Diane
; REGISTRATION NUMBER: 40,943
; REFERENCE/DOCKET NUMBER: 4952, US. D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847-938-3137
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3567 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-439-009A-4

Query Match 55.4%; Score 884.5; DB 3; Length 3567;
Best Local Similarity 57.3%; Pred. No. 2.6e-80;
Matches 177; Conservative 41; Mismatches 88; Indels 3; Gaps 2;
QY 1 VFVPPGGGGMAGNGVRLASSPVFARRMQACEBALAPWVDSVVDILRRDAGDAVERA 60
Db 560 VFLEPGGSGMAGNGAELLSSPVFAGKIRACDSMA PMQDWKXSVLRRQAPAGLDIV 619
QY 61 DVVQPVLFSTWVSLAALMRSYGIIPDAVLGHSGCEIAAAHVCGALSLKDAKTVALSRA 120
Db 620 DVVQPVLFSTWVSLAALMRSYGIIPDAVLGHSGCEIAAAHVCGALSLKDAKTVALSRA 120
QY 121 LAAVRGGGMAVPLPAOEVEQLIGERWAGRLWVAANGPRSTAVSGDAEAVDEVLA YCA 180
Db 680 MRSLSGGGMAVPLPAOEVEQLIGERWAGRLWVAANGPRSTAVSGDAEAVDEVLA YCA 180
QY 181 GTGVARRIRPVYASHCNVPOLREBELLELIGDISPPSGVFPFSTYEGTWLDTTITDAA 240
Db 739 AEGIRVADIDVDYASHSPQIERVREBELLETTGDIAPRPARVTFHSTVESRMDGTELDAR 798
QY 241 YWYRNHLQPVRFSDAVOALADGHRVFEVSPHPTLVPAIEDTTE--DTAEDVTAIGSLR 298
Db 799 YWYRNHLRETVRFADAVTRLAESGYDAFIEVSPHVVVQAVEAEVADGAEADAVVVGSLH 858
QY 299 RGNDRTRF 307

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-4

Query Match 53.1%; Score 848.5; DB 4; Length 3170;
Best Local Similarity 56.2%; Pred. No. 1e-76;

Matches 172; Conservative 35; Mismatches 98; Indels 1; Gaps 1;

QY 1 VFVFPQGGQWGMAGVRLTASSPVFARMOACEALAPWVDSVVDILRRDAGDAWERA 60
DB 563 VFVFPQGGQWGMAGVRLTASSPVFARMOACEALAPWVDSVVDILRRDAGDAWERA 622
QY 61 DVVQPVLFVSVVSLAALMRSYGIETPAVIGHSOGSEIAAAHVCGALSLKDAKTVALSRA 120
DB 623 DVVQPVLFVSVVSLAALMRSYGIETPAVIGHSOGSEIAAAHVCGALSLKDAKTVALSRA 682
QY 121 LAAVRRGGMAVPLPAQVEBQILGERMAGRLVVAANGRSTAVSGDAEVADEVLAYCA 180
DB 683 LKGLSGRGMAVSLACPADEVAALFAGS-GRLEVAALINGRSTAVSGDAEVADELLAECA 741
QY 181 GTGVRARRIPVDYASHCPHVPQPLREBELLELDISPQSGVPFSTVEGTWLDPTTLDDAA 240
DB 742 EKDMRRRIPVDYASHSAHVEVRSVPLAAAGVRHDDGVPMWSTVIGWVDPARLDGE 801
QY 241 WYVRNLHQVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDPAEDVTAIGSLRG 300
DB 802 WYVRNLHQVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDPAEDVTAIGSLRG 861
QY 301 DNDTR 306
DB 862 SGGLR 867

RESULT 9

US-09-370-700-4
Sequence 4, Application US/09370700
Patent No. 6274350
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patci J
APPLICANT: Turner, Jan R
TITLE OF INVENTION: Biochemical Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 DIV1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 4
LENGTH: 3170
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-4

Query Match 53.1%; Score 848.5; DB 4; Length 3170;
Best Local Similarity 56.2%; Pred. No. 1e-76;
Matches 172; Conservative 35; Mismatches 98; Indels 1; Gaps 1;

QY 1 VFVFPQGGQWGMAGVRLTASSPVFARMOACEALAPWVDSVVDILRRDAGDAWERA 60
DB 563 VFVFPQGGQWGMAGVRLTASSPVFARMOACEALAPWVDSVVDILRRDAGDAWERA 622
QY 61 DVVQPVLFVSVVSLAALMRSYGIETPAVIGHSOGSEIAAAHVCGALSLKDAKTVALSRA 120
DB 623 DVVQPVLFVSVVSLAALMRSYGIETPAVIGHSOGSEIAAAHVCGALSLKDAKTVALSRA 682
QY 121 LAAVRRGGMAVPLPAQVEBQILGERMAGRLVVAANGRSTAVSGDAEVADEVLAYCA 180
DB 683 LKGLSGRGMAVSLACPADEVAALFAGS-GRLEVAALINGRSTAVSGDAEVADELLAECA 741

DB 683 LKGLSGRGMAVSLACPADEVAALFAGS-GRLEVAALINGRSTAVSGDAEVADELLAECA 741
QY 181 GTGVRARRIPVDYASHCPHVPQPLREBELLELDISPQSGVPFSTVEGTWLDPTTLDDAA 240
DB 742 EKDMRRRIPVDYASHSAHVEVRSVPLAAAGVRHDDGVPMWSTVIGWVDPARLDGE 801
QY 241 WYVRNLHQVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDPAEDVTAIGSLRG 300
DB 802 WYVRNLHQVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDPAEDVTAIGSLRG 861
QY 301 DNDTR 306
DB 862 SGGLR 867

RESULT 10

US-07-642-734C-2
Sequence 2, Application US/07642734C
Patent No. 5824513
GENERAL INFORMATION:
APPLICANT: Donadio, S
APPLICANT: Katz, L
APPLICANT: McAlpine, J B
TITLE OF INVENTION: Recombinant DNA Method for Producing
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Edward H. Gorman
STREET: Abbott Laboratories D377/AP6D-2 One Abbott
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/642,734C
FILING DATE: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Danckert, Andreas M
REGISTRATION NUMBER: 32652
REFERENCE/DOCKET NUMBER: 4952.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-642-734C-2

Query Match 52.6%; Score 840.5; DB 2; Length 3491;
Best Local Similarity 55.1%; Pred. No. 7.7e-76;
Matches 173; Conservative 46; Mismatches 84; Indels 11; Gaps 4;

QY 1 VFVFPQGGQWGMAGVRLTASSPVFARMOACEALAPWVDSVVDILRRDAGDAWERA 60
DB 1031 VFVFPQGGQWGMAGVRLTASSPVFARMOACEALAPWVDSVVDILRRDAGDAWERA 1090
QY 57 --WERADVQPVLFVSVVSLAALMRSYGIETPAVIGHSOGSEIAAAHVCGALSLKDAKTV 114
DB 1091 LSTERADVQPVLFVSVVSLAALMRSYGIETPAVIGHSOGSEIAAAHVCGALSLKDAKTV 1150
QY 115 ALRSRALAAVRRGGMAVPLPAQVEBQILGERMAGRLVVAANGRSTAVSGDAEVADE 174
DB 1151 ALRSRALAAVRRGGMAVPLPAQVEBQILGERMAGRLVVAANGRSTAVSGDAEVADE 1206

OY 121 LAAVRGGGMAVPELPAQVEEOLIGRMWAGRLVVAVNGRSTAVSGDAEAVDEVLAYCA 180
125 WLGLAGGGMVAVPMPEBELRPL-VTWGRDLAAVAVNSPGSCAVAGDPALAEVLALT 183
OY 181 GTGVRARIP-VDYASHCPVOPRLREBELLDGDISFQPSGVPPFSTVEGTWLDTTLLDA 239
184 GEGVHARPIPGVDTAGHSPQVDALRAHLEVLAPVAPRPADI PFYSTVTGGLDGTBLDA 243
OY 240 AYYRNKHOPVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETEDTAEDVTAIGSLR 299
244 TYWRNRREPEFERATRALIADGHDFLETSHPMLAVALLEQVTVDAGTDAVAGTLRR 303

RESULT 13

US-09-078-166-33
Sequence 33, Application US/09078166
Patent No. 6063561
GENERAL INFORMATION:
APPLICANT: Katz, Leonard
APPLICANT: Staszi, Diane L.
APPLICANT: Summers Jr., Richard G.
APPLICANT: Ruan, Xiaolan
APPLICANT: Pereda-Lopez, Ana
TITLE OF INVENTION: NOVEL POLYKETIDE DERIVATIVES
TITLE OF INVENTION: AND RECOMBINANT METHODS FOR MAKING SAME
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Rd.
CITY: Abbott Park
STATE: Illinois
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/078,166
FILING DATE: 16-MAY-1979
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dianne Casuto
REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 4952.US.P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (847)-938-3137
TELEFAX: (847)-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6063561
US-09-078-166-33

Query Match 51.8%; Score 828; DB 3; Length 344;
Best Local Similarity 55.0%; Pred. No. 4,3e-76;

Matches 165; Conservative 41; Mismatches 92; Indels 2; Gaps 2;

OY 1 VFVFGGQGMAGVRLASSPVFARRMOACEALAPWDMGVVDILRRDAGDAVAVERA 60
DB 5 VFVFGGQGMAGVRLASSPVFARRMOACEALAPWDMGVVDILRRDAGDAVAVERA 64
OY 61 DVVQPVLFVWVSLAALMRSGYIEPDAVLGHSGEIAAAHVCGALSLKDAKTVALSRA 120

DB 65 DVVQPVLFVWVSLAALMRSGYIEPDAVLGHSGEIAAAHVCGALSLKDAKTVALSRA 124
OY 121 LAAVRGGGMAVPELPAQVEEOLIGRMWAGRLVVAVNGRSTAVSGDAEAVDEVLAYCA 180
125 WLGLAGGGMVAVPMPEBELRPL-VTWGRDLAAVAVNSPGSCAVAGDPALAEVLALT 183
OY 181 GTGVRARIP-VDYASHCPVOPRLREBELLDGDISFQPSGVPPFSTVEGTWLDTTLLDA 239
184 GEGVHARPIPGVDTAGHSPQVDALRAHLEVLAPVAPRPADI PFYSTVTGGLDGTBLDA 243
OY 240 AYYRNKHOPVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETEDTAEDVTAIGSLR 299
244 TYWRNRREPEFERATRALIADGHDFLETSHPMLAVALLEQVTVDAGTDAVAGTLRR 303

RESULT 14

US-08-997-467-33
Sequence 33, Application US/08997467
Patent No. 6200813
GENERAL INFORMATION:
APPLICANT: Katz, Leonard
APPLICANT: Staszi, Diane L.
APPLICANT: Summers Jr., Richard G.
APPLICANT: Ruan, Xiaolan
APPLICANT: Pereda-Lopez, Ana
TITLE OF INVENTION: NOVEL POLYKETIDE DERIVATIVES
TITLE OF INVENTION: AND RECOMBINANT METHODS FOR MAKING SAME
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Rd.
CITY: Abbott Park
STATE: Illinois
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,467
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/858,003
FILING DATE: 16-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dianne Casuto
REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 4952.US.P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (847)-938-3137
TELEFAX: (847)-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6200813
US-08-997-467-33

Query Match 51.8%; Score 828; DB 4; Length 344;
Best Local Similarity 55.0%; Pred. No. 4,3e-76;

Matches 165; Conservative 41; Mismatches 92; Indels 2; Gaps 2;

OY 1 VFVFGGQGMAGVRLASSPVFARRMOACEALAPWDMGVVDILRRDAGDAVAVERA 60
DB 5 VFVFGGQGMAGVRLASSPVFARRMOACEALAPWDMGVVDILRRDAGDAVAVERA 64

Qy	61	UVVGVGLSVVWVSLAALMSYGIIEPRAVYGHSGEITAAAHVCGALSLKQAAKVTVALRSR	120
Db	65	DVVGCVLTLTMVSLAAVMRALGVEPAAVGHSGEITAAAHVCGALSLDSDARVVALRSR	124
Qy	121	LAAYRGGMASVPLPAQEVQLIGERWAGRLVAAVNPGRSTAVSGDAEVADEVIAVCA	180
Db	125	WLGTLGKGMGVAVPMPAEELRPRL-ITWGDRLAAVAVNSPGSCAVGDEPEALAEVALLT	183
Qy	181	GTGVARARIRP-VDYASHCHVQPLRELLLELLGDISIPQSGVFFSTVEGTLDDTTLDA	239
Db	184	GEGVARARIRPGVDTGHSPOVALRALHLLLEVLAPVAPRADIPFIYETVIVGGLLDGTELLA	243
Qy	240	AYWYENLQPVFESDAVOLADDDGHRVFEVSPHPTLVPAIEDTETDTEADVTAIGSLRR	299
Db	244	TYWYNNMEPVEFEAKTRALLADGHDVFELETSPHPLVALLEDVTVDAGTDAAVLGTLRR	303

```

RESULT 15
US-08-804-227C-10
Sequence 10, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhse, Stuart A.
APPLICANT: Roseck, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYMERIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-376-2459
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 3724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-227C-10

```

Query Match	51.5%;	Score 822.5;	DB 2;	Length 3724;
Best Local Similarity	55.0%;	Pred. No. 5.7e-74;		
Matches 169;	Conservative 36;	Mismatches 97;	Indels 5;	Gaps 2;

Qy	2	FVFPQGGOMGMGRVILLASPVFAARMQACEALAPWUMDSVITLREDDAGDAVWERAD	61
Db	594	FVFPQGGOMGMGRVILLDTAFAFAEEDRCGALS PYVDMNADVLKGPAPAPGLDRAD	653
Qy	62	WVQVLPFSVMSLAALWRSYGIIEPDVAVLGHSGOEIAAAHVCGLSLKDAKTVALSRAI	121
Db	654	WVQVLPFAVMVGLAALWRSLSGVEPAVIVIGHSGOEIAAACVAGALSLEDAAIRIVALSQYI	713
Qy	122	A-AVNRGRGMSVPLPAQVEEQLIGERNMGRIMVAANVNPRTSTASGDAEVADEVATAYCA	180

```

Db      71:  ARLAAGRGMAVALPAAEVE----ARLAGVEIAAANVGSTVACEPGALAEALVTLE 769

Qy      181  GTCVARRRIPVDYASHCHPVCPRLRELLIGDISPODSGVPEFSTVEGWLDTTTIDA 240
        SESTRVRRIDVVYASHSHVESIRAEILATVIGAPRPRGDPVFYSTVEALLDPATLAD 829

Db      770  SESTRVRRIDVVYASHSHVESIRAEILATVIGAPRPRGDPVFYSTVEALLDPATLAD 829

Qy      241  YWVRNRLHGVRRPSDAVQALADGHRVFEVSPHPTLVPAIEDTEADYAEVTAIGSLRG 300
        ||||| ||||| :::: ||| ||||| ||||| :::: ||||| ||||| 830
        YWVRNRLHGVRRPEPTVRAALDDGVDAFVACSAPHVLTIVGQTVESAGAVPALASLRD 889

Db      830  YWVRNRLHGVRRPEPTVRAALDDGVDAFVACSAPHVLTIVGQTVESAGAVPALASLRD 889

Qy      301  DNDTRRF 307
        : |||
Db      890  EGGIRRF 896

```

Search completed: June 17, 2003, 13:09:26
Job time : 6.57989 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 12:50:57 / Search time 15.0391 Seconds
(without alignments)
2720.102 Million cell updates/sec

Title: US-09-914-286-3_COPY_1050_1356

Perfect score: 1598
Sequence: 1 VFVFPQGGCGMAGVRLA.....AEDVTAIGSHRQNDNTRRF 307

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1598	100.0	3972	21	AA823749
2	1598	100.0	3972	22	AA655264
3	1598	100.0	3972	22	AA655268
4	1598	100.0	5532	21	AA823752
5	1598	100.0	5532	22	AA655267
6	1366	85.5	6239	21	AA833750
7	1366	85.5	6239	22	AA655265
8	891.5	55.8	3816	21	AA92708
9	884.5	55.4	3567	14	AA44431
10	861.5	53.9	4150	21	AA92707

11	858	53.7	3519	21	AA92709	S. antibioticus 8,
12	852	53.3	2595	20	AA39297	Spn a polyketide
13	852	53.3	2595	22	AA70965	S. spinosa protein
14	848.5	53.1	3170	20	AA73929	Spn a polyketide
15	848.5	53.1	3170	22	AA70967	S. spinosa protein
16	848.5	53.1	3069	19	AA52846	A. mediterranei xl
17	840.5	52.6	3398	14	AA44430	eryA region polype
18	832	52.1	3546	22	AA84432	Polyketide synthase
19	832	52.1	4881	21	AA823751	S. avermitilis ave
20	832	52.1	4881	22	AA65266	Streptomycetes avem
21	828.5	51.8	3192	22	AA810128	Streptomycetes avem
22	828	51.8	344	20	AA87714	An active acyltran
23	828	51.8	344	21	AA812926	Protein sequence O
24	827	51.8	3562	22	AA882213	Polyketide synthase
25	826	51.7	7068	22	AA810142	Streptomycetes nous
26	825	51.7	9477	22	AA810144	Streptomycetes nous
27	822.5	51.5	3724	18	AA823718	Platenolide synthase
28	822.5	51.5	3724	18	AA82608	Platenolide synthase
29	815.5	51.0	4572	19	AA52845	A. mediterranei xl
30	814.5	51.0	2986	14	AA44432	eryA region polype
31	812.5	50.8	1721	19	AA52847	A. mediterranei xl
32	809.5	50.7	3201	22	AA882214	Polyketide synthase
33	806.5	50.5	3413	19	AA52849	A. mediterranei xl
34	805	50.4	1688	19	AA52848	A. mediterranei xl
35	774	48.4	5588	20	AA739301	Spn a polyketide
36	774	48.4	5588	22	AA870969	S. spinosa protein
37	773	48.4	1580	18	AA823719	Platenolide synthase
38	773	48.4	1580	18	AA82609	Platenolide synthase
39	772.5	48.3	1346	21	AA818640	Amino acid sequenc
40	772.5	48.3	1346	21	AA77195	A. venezuelae macr
41	772.5	48.3	1346	21	AA67204	Narboxolide synthase
42	772.5	48.3	3739	21	AA818638	Amino acid sequenc
43	772.5	48.3	3739	21	AA772193	S. venezuelae macr
44	772.5	48.3	3739	21	AA772193	S. venezuelae pik
45	772.5	48.3	3739	21	AA67202	Narboxolide synthase

ALIGNMENTS

RESULT 1	
AA823749	AA823749 standard; Protein; 3972 AA.
ID	AA823749
XX	AA823749;
AC	10-JAN-2001 (first entry)
DT	
XX	
DE	S. avermitilis avermectin aglycon synthase protein SEQ ID NO:3.
XX	
KW	Streptomycetes avermitilis; avermectin aglycon synthase; biosynthesis;
KW	multifunctional enzyme; polyketide; avermectin; veterinary drug;
KW	agrochemical.
XX	
OS	Streptomycetes avermitilis.
XX	
PN	W0200050605-A1.
XX	
PD	31-AUG-2000.
XX	
PF	23-FEB-2000; 2000MO-JD01041.
XX	
PR	24-FEB-1999; 99JP-0046961.
XX	
PA	(KITA) KITASATO INST.
XX	
PI	Omura S, Ikeda H;
XX	
DR	WPI: 2000-565458/52.
XX	
DR	N-PSDB; AAA92301.
XX	
PT	Avermectin aglycone synthase DNA and proteins encoded by all or part of
	it for the production of avermectin and its derivatives for drug and

PT agrochemical use -
 XX
 PS Claim 32; Page 203-223; 314pp; Japanese.
 CC
 CC The present sequence represents an avermectin aglycon synthase protein.
 CC Also described are: (1) polypeptides encoded by all or part of the DNA;
 CC (2) expression vectors containing the DNA; (3) host cells transformed by
 CC the vectors; (4) preparation of the polypeptides by culture of the
 CC transformants; (5) preparation of avermectin aglycon or its derivatives
 CC by culture of transformed avermectin-producing microorganisms; and (6)
 CC oligonucleotides of 5-60 bases in length containing sense or antisense
 CC sequences from the avermectin aglycon synthase DNA. The enzymes are
 CC useful for the production of modified forms of avermectin and of the
 CC intermediates in its biosynthesis, for use as drugs, veterinary drugs
 CC and agrochemicals.
 XX
 SQ Sequence 3972 AA;
 Query Match 100.0%; Score 1598; DB 21; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 2.1e-148;
 Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VFVPPGGGGWAGMGVRLASSPVFARRMQACEBALPWTWMSVVDILRRDAGDAWVERA 60
 DB 1050 VFVPPGGGGWAGMGVRLASSPVFARRMQACEBALPWTWMSVVDILRRDAGDAWVERA 1109
 QY 61 DVVQPVLFPSVWVSLAALMRSYGIEPDVILGHSGEIAAAHVCGALSLKDAKTVALSRA 120
 DB 1110 DVVQPVLFPSVWVSLAALMRSYGIEPDVILGHSGEIAAAHVCGALSLKDAKTVALSRA 1169
 QY 121 LAAVRGGGMAVPLPAQVEQLIGERWAGRLWVAANGPRSTAVSGDAEAVDEVLAYCA 180
 DB 1170 LAAVRGGGMAVPLPAQVEQLIGERWAGRLWVAANGPRSTAVSGDAEAVDEVLAYCA 1229
 QY 181 GTGVRARRIPVDYASHCPHQPRLREELIELLGDISPQSGVPFSTVEGTLDTTLLDAA 240
 DB 1230 GTGVRARRIPVDYASHCPHQPRLREELIELLGDISPQSGVPFSTVEGTLDTTLLDAA 1289
 QY 241 YWYRNLIHQPVRFSDAVQALADDGHRVFEVSPHPTLPAIEDTETDAEDVTAIGSLRIG 300
 DB 1290 YWYRNLIHQPVRFSDAVQALADDGHRVFEVSPHPTLPAIEDTETDAEDVTAIGSLRIG 1349
 QY 301 DNDTRRF 307
 DB 1350 DNDTRRF 1356
 RESULT 2
 AAG65264
 ID AAG65264 standard; Protein; 3972 AA.
 XX
 AC AAG65264;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Streptomycetes avermilitilis protein SEQ ID NO: 4.
 XX
 KM Avermectin aglycone synthase; AAS; avermectin derivative;
 KM drug production; veterinary drug; pesticide.
 XX
 OS Streptomycetes avermilitilis.
 XX
 PN WO200162939-A1.
 XX
 PD 30-AUG-2001.
 XX
 PF 23-FEB-2001; 2001WO-JP01381.
 XX
 PR 24-FEB-2000; 2000JP-0047405.
 XX
 PA (KYOWA) KYOWA HAKKO KOGYO KK.
 PA (KITA) KITASATO INST.
 XX

PI Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;
 XX
 DR WPI; 2001-582053/65.
 DR N-PSDB; AAH79277.
 XX
 PT New modified avermectin aglycone synthase derived from Streptomycetes.
 PT avermectin used in production of 22,23-dihydroavermectin B1a used in
 PT drugs and pesticides -
 XX
 PS Claim 4; Page 167-180; 257pp; Japanese.
 XX
 CC The present invention relates to the production of modified derivatives
 CC of avermectin aglycone synthase (AAS) derived from Streptomycetes
 CC avermilitilis. The activity of an acyl carrier protein (ACP).
 CC beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT),
 CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl
 CC reductase (ER) and/or thioesterase (TE) domain may be reduced or
 CC suppressed. The process can be used in the production of drugs, veterinary
 CC drugs and pesticides. The present sequence is an S. avermilitilis protein.
 XX
 SQ Sequence 3972 AA;
 Query Match 100.0%; Score 1598; DB 22; Length 3972;
 Best Local Similarity 100.0%; Pred. No. 2.1e-148;
 Matches 307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VFVPPGGGGWAGMGVRLASSPVFARRMQACEBALPWTWMSVVDILRRDAGDAWVERA 60
 DB 1050 VFVPPGGGGWAGMGVRLASSPVFARRMQACEBALPWTWMSVVDILRRDAGDAWVERA 1109
 QY 61 DVVQPVLFPSVWVSLAALMRSYGIEPDVILGHSGEIAAAHVCGALSLKDAKTVALSRA 120
 DB 1110 DVVQPVLFPSVWVSLAALMRSYGIEPDVILGHSGEIAAAHVCGALSLKDAKTVALSRA 1169
 QY 121 LAAVRGGGMAVPLPAQVEQLIGERWAGRLWVAANGPRSTAVSGDAEAVDEVLAYCA 180
 DB 1170 LAAVRGGGMAVPLPAQVEQLIGERWAGRLWVAANGPRSTAVSGDAEAVDEVLAYCA 1229
 QY 181 GTGVRARRIPVDYASHCPHQPRLREELIELLGDISPQSGVPFSTVEGTLDTTLLDAA 240
 DB 1230 GTGVRARRIPVDYASHCPHQPRLREELIELLGDISPQSGVPFSTVEGTLDTTLLDAA 1289
 QY 241 YWYRNLIHQPVRFSDAVQALADDGHRVFEVSPHPTLPAIEDTETDAEDVTAIGSLRIG 300
 DB 1290 YWYRNLIHQPVRFSDAVQALADDGHRVFEVSPHPTLPAIEDTETDAEDVTAIGSLRIG 1349
 QY 301 DNDTRRF 307
 DB 1350 DNDTRRF 1356
 RESULT 3
 AAG65268
 ID AAG65268 standard; Protein; 3972 AA.
 XX
 AC AAG65268;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Streptomycetes avermilitilis protein derivative SEQ ID NO: 8.
 XX
 KM Avermectin aglycone synthase; AAS; avermectin derivative;
 KM drug production; veterinary drug; pesticide.
 XX
 OS Synthetic.
 XX
 PN WO200162939-A1.
 XX
 PD 30-AUG-2001.
 XX
 PF 23-FEB-2001; 2001WO-JP01381.
 XX
 PR 24-FEB-2000; 2000JP-0047405.
 XX

KW	Avermectin aglycone synthase; AAs; avermectin derivative;
KV	drug production; veterinary drug; pesticide.
XX	
OS	Streptomyces avermiltis.
PN	WO200162939-A1.
PD	30-AUG-2001.
XX	
PF	23-FEB-2001; 2001WO-JP01381.
PR	24-FEB-2000; 2000JP-0047405.
XX	
PA	(KYOM) KYOMA HAKKO KOGYO KK. (KITA) KITASATO INST.
XX	
PI	Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H; WPI; 2001-582053/65.
DR	N-PADB; AAH79278.
XX	
PT	New modified avermectin aglycone synthase derived from Streptomyces avermecilis used in production of 22,23-dihydroavermectin B _{1a} used in drugs and pesticides -
PS	Claim 4; Page 217-235; 257pp; Japanese.
CC	The present invention relates to the production of modified derivatives of avermectin aglycone synthase (AAS) derived from Streptomyces avermicilli. The activity of an acyl carrier protein (ACP), beta-ketoacyl carrier protein synthase (KS), acyltransferase (AT), CC beta-ketoacyl carrier protein reductase (KR), dehydratase (DH), enoyl reductase (ER) and/or thioesterase (TE) domain may be reduced or suppressed. The process can be used in the production of drugs, veterinary drugs and pesticides. The present sequence is an S. avermiltis protein.
SQ	Sequence 5532 AA;
Query Match	100.0%; Score 1598; DB 22; Length 5532;
Best Local Similarity	100.0%; Pred. No. 3.5e-148;
Matches 307; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Db	1 VVFVFGGCGQAAGMGVRLLASSPVFARRRQACEEALAEVMDSVDILRRDAGAAMWRA 60
Qy	VVFVFGGCGQAAGMGVRLLASSPVFARRRQACEEALAEVMDSVDILRRDAGAAMWRA 60
Db	603 VFVFFGGGQMAAGMGRRLASSPVFAARRRQACEEALAEMVSVDILRRDAGAAMWERA 622
Qy	DVVQGVLFSVMWSLAALMRYSYIEEPDAVLGHSGGEIAAAHVGCALSGLDKAAKTVALRSRA 120
Db	663 DVVGQVLFSVMWSLAALMRYSYIEEPDAVLGHSGGSIAAHVCAGLSLDDAKTVALRSRA 722
Qy	121 LAAVGRGGMASVPLPAOEVEQLIGERNAGRLVAANVPSTAVSGAEADEVLAYCA 180
Db	723 LAAVGRGGMASVPLPAOEVEQLIERVWAGRMLVAANVGPRSTAUSGDAAEVDEVLAYCA 782
Qy	181 GGCVARRRIPRVUYASHCHVOPFLREBELLELDGISPOPSGVFPFTVGTWLDTTLDDAA 240
Db	783 GGCVARRRIPRVUYASHCHVOPFLREELLELDISPOPSGVFPFTVGTWLDTTLDDAA 842
Qy	241 WYRNLIHQVFRSDAIVQALLADGGHVFVEVSEPHPTLVPAIEDTTEDTDVEDYTAIGSLRG 300
Db	843 WYRNLIHQVFRSDAIVQALLADGGHVFVEVSHPLTVPAIETTEDTDVEDYTAIGSLRG 902
Qy	301 DNDTRRF 307
Db	903 DNDTRRF 909
RESULT 6	
AAB23750	
ID	AAB23750 standard; Protein; 6239 AA.
AC	AAB23750;
XX	

[illegible]

ID	AA655265	standard; Protein; 6239 AA.
AC	AA655265;	
DT	04-DEC-2001	(first entry)
DE	Streptomyces avermiltillis protein SEQ ID NO: 5.	
KM	Avermectin aglycone synthase; AAs, avermectin derivative; drug production; veterinary drug; pesticide.	
OS	Streptomyces avermiltillis.	
PN	MO200162939-AA.	
PD	30-AUG-2001.	
PP	23-FEB-2001; 2001MO-JP01381.	
PR	24-FEB-2000; 2000JP-0047405.	
PA	(KYOW) KYOWA HAKKO KOGYO KK. (KITA) KITABATO INST.	
P1	Endo H, Yamaguchi H, Kanda Y, Hashimoto S, Omura S, Ikeda H;	
DR	MPJ; 2001-582053/65.	
DR	N-P8DB; AAH79277.	
PT	New modified avermectin aglycone synthase derived from Streptomyces avermiltillis used in production of 22,23-dihydroavermectin B1a used in drugs and pesticides -	
PS	Claim 4; Page 180-201; 257pp; Japanese.	
XX	The present invention relates to the production of modified derivatives of avermectin aglycone synthase (AAS) derived from Streptomyces CC CC avermiltillis. The activity of an acyl carrier protein (ACP), CC beta-ketocacyl carrier protein synthase (KS), acyltransferase (AT), CC beta-ketocacyl carrier protein reductase (KR), dehydratase (DH), enoyl CC reductase (ER) and/or thioesterase (TE) domain may be reduced or suppressed. The process can be used in the production of drugs, veterinary CC drugs and pesticides. The present sequence is an S. avermiltillis protein. CC	
XX	Sequence 6239 AA;	
XX		
Query Match	85.5%; Score 1366; DB 22; Length 6239;	
Beet Local Similarity	81.8%; Pred. No. 4.3e-125;	
Matches 251; Conservative 25; Mismatches 31; Indels 0; Gaps 0;		
Ds	1 VFVFGGCGAGNAGVRLIASSPVFARMOACEALAPMVMSVVDILRRDAGDAVMERA 60 4815 VFVFPGGGCGNAGMGTLTSSPFVAEHIIDACEKALTPTWVMSLTDILHRPDDPMQQA 4874	
Oy	61 DVQVPRLFVSIVSLAALMRSTGIIEPDVVLGHSGGEIAAHAVCGALSLKDAKTVALRSRA 120 4875 DVQVPRLFVSIMWSLAALMRSTGIIEPDVVLGHSGGEIAAHAVCGALSLKDAKTVALRSRA 4934	
Oy	121 LAAYARGRGMAVPLPAOEVBQLGERWAGRMLVAANVGPRSTANGSGLAEVDEVLYACA 180 4935 LAAYARGRGMAVPLPAODVQQLISEREGQLVAALNGPHSTTVSGDTKANDEVYLAHCT 4994	
Oy	181 GTGVARRIPVDYASHCHVOPFLREELLELLGDISPQSGLVPFSEGTVEGWLDTITLDAA 240 4995 DTGLAKRIRPDYASHCHVOPFLREELLELLGDITPOPSTVFPFSEGTVEGWLDTITLDAA 5054	
Oy	241 TWYRNLHPVRPVSADVOALLDGHHVPEVSHPTLVPAIEDTETEDTADTAAGSLARG 300 5055 TWYRNLHPVRPVSNAIQTLTDGHAAPLEISHPTLVPAIEDTENTENTENTATGSLARG 5114	

QY	301	DNDTRF	307
DB	5115	DNDTRF	5121
RESULT 8			
AA92708			
ID	AA92708	standard; Protein; 3816 AA.	
AC	AA92708;		
XX			
DT	29-AUG-2000	(first entry)	
XX			
DE	S. antibioticus 8,8a-deoxyoleandolide synthase II encoded by oleAII.		
XX			
KW	oleandomycin; oleandolide; polyketide synthase; oleAI; oleAII; oleAIII		
KW	PKS; type I; 8,8a-deoxyoleandolide synthase; modular; ketosynthase;		
KW	acyl-transferase; acyl carrier protein; inactivated; polyketide;		
KW	macrolactone; antibiotic; mactilide; erythromycin.		
XX			
OS	Streptococcus antibioticus.		
XX			
PH	Key	Location/Qualifiers	
FT	Region	30..1573	
FT		/label= Extender_module_3	
FT	Domain	30..459	
FT		/label= KS3_domain	
FT		/note= "ketosynthase domain"	
FT	Domain	566..911	
FT		/label= AT3 domain	
FT		/note= "acyl transferase domain"	
FT	Domain	1210..1394	
FT		/label= KR3_domain	
FT		/note= "ketoreductase domain"	
FT	Domain	1487..1573	
FT		/label= ACP3	
FT		/label= ACP3 domain	
FT	Region	1593..3710	
FT		/note= "acyl carrier protein domain"	
FT	Domain	1593..3710	
FT		/label= Extender_module_4	
FT		1593..2021	
FT	Domain	/label= KS4 domain	
FT		/note= "ketosynthase domain"	
FT	Domain	2126..2472	
FT		/label= AT4_domain	
FT		/note= "acyl transferase domain"	
FT	Domain	2484..2663	
FT		/label= DH4 domain	
FT		/note= "dehydratase domain"	
FT	Domain	3054..3345	
FT		/label= ER4_domain	
FT		/note= "enoyl reductase domain"	
FT	Domain	3349..3532	
FT		/label= KR4 domain	
FT		/note= "ketoreductase domain"	
FT	Domain	3627..3710	
FT		/label= ACP4 domain	
FT		/note= "acyl carrier protein domain"	
PV	MO200026349-A2.		
XX			
PD	11-MAY-2000.		
PF	22-OCT-1999;	99WO-US24478.	
XX			
PR	29-OCT-1998;	98US-0106100.	
PR	16-FEB-1999;	99US-0120254.	
XX			
PA	(KOSA-) KOSAN BIOSCIENCES INC.		
XX			
PI	Beclach WC, Shah SK, McDaniel R, Tang L;		
XX			
DR	WPI; 2000-365602/31.		
DR	N-PSDB; AAA09469.		

XX Recombinant DNA compound encoding oleandrolide polyketide synthase for
PT synthesizing polyketides comprising a coding sequence for a domain of a
PT loading module or any one of extender modules

PS Disclosure, Page 28-29; 86pp; English.

XX The oleandrolide polyketide synthase (PKS), also known as
CC 8,8a-deoxyoleandrolide synthase, is encoded by three open reading frames
CC (ORF), designated oleaI, oleaII and oleaIII. The PKS is a type I
CC "modular" enzyme, where each ORF encodes 2 extender modules and
CC the first ORF also encodes the loading module. Each module is composed
CC of at least a ketosynthase (KS), acyl-transferase (AT) and an
CC acyl carrier protein (ACP) domain. The oleandrolide PKS loading module
CC contains an inactivated KS, called KS-Q, where Q is the abbreviation for
CC glutamine, present instead of the active site cysteine required for
CC activity. The large multifunctional PKS enzymes catalyze the biosynthesis
CC of polyketide macroketones through multistep pathways involving
CC decarboxylative condensations between acylthioesters followed by cycles
CC of varying beta-carbon processing activities. The macroide product of
CC the PKS, 8,8a-deoxyoleandrolide, is further modified by epoxidation and
CC glycosylation to yield oleandromycin, an antibacterial polyketide. The
CC invention concerns an isolated recombinant DNA compound, comprising a
CC coding sequence for a domain of loading module or any one of extender
CC modules 1-4 or 1-6, including an oleandrolide PKS operably linked to a
CC promoter. Also discussed are recombinant oleandrolide PKS in which the
CC module 1 KS domain is inactivated by deletion or other mutation. In
CC particular, the inactivation is mediated by a change in the KS domain
CC that renders it incapable of binding substrate (the KS1-o mutation),
CC rendered by mutation in the codon for the active site cysteine. The
CC oleandrolide PKS is useful for synthesizing polyketides, which are useful
CC as antibiotics and molluscs. Heterologous expression of oleandrolide PKS
CC in host cells such as Streptomyces coelicolor and S. lividans is also
CC made possible. Unmodified oleandrolide compounds can be provided to the
CC cultures of Saccharopolyspora erythraea and converted to the
CC corresponding derivatives of erythromycins A-D.

SQ Sequence 3816 AA;

Query Match 55.8%; Score 891.5; DB 21; Length 3816;

Best Local Similarity 58.6%; Pred. No. 2.6e-76; Matches 180; Conservative 31; Mismatches 93; Indels 3; Gaps 2;

QY 1 VFVPPGCGGAGWGVRLASSPVFARRMQACEBALAPWVDSVVDILRRDAGDAWERA 60
DB 567 VFVPPGCGGAGWGVRLASSPVFARRMQACEBALAPWVDSVVDILRRDAGDAWERA 60
QY 61 DVVQPVLFSSVMSIALMRSYGIPEPAVLGHSGEIAAAHVCGALSLKDAKTVALLRSRA 120
DB 625 DVVQPVLFSSVMSIALMRSYGIPEPAVLGHSGEIAAAHVCGALSLKDAKTVALLRSRA 120
QY 121 LAAVRGRGMAVSLPDAOEVEQLIGERMAGRLVAAVNGPRSTVSGDAEAVDEVLYACA 180
DB 685 LAAVRGRGMAVSLPDAOEVEQLIGERMAGRLVAAVNGPRSTVSGDAEAVDEVLYACA 180
QY 181 GTGVARRIPVDYASHCPHVPOLREBELLELGDISPQSGVPPFSTVEGTWLDTTTLDA 240
DB 744 REGRRARIPVDYASHCPHVPOLREBELLELGDISPQSGVPPFSTVEGTWLDTTTLDA 240
QY 241 YWYRNLIHQVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 300
DB 804 YWYRNLIHQVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 300
QY 301 DNDTRRF 307
DB 864 EGGIGRF 870

RESULT 9
AAR44431
ID AAR44431 standard; Protein; 3567 AA.
XX
AC AAR44431;

XX 22-DEC-1993 (first entry)
DT
XX
PT 22-DEC-1993 (first entry)

DE 22-DEC-1993 (first entry)

XX Saccaropolyspora erythraea; eryA; biosynthesis; polyketide; module;
KW erythromycin; condensation; elongation; acyl chain growth;
KM gene replacement.

OS Saccharopolyspora erythraea.
XX
XX WO9313663-A.
PN
XX 22-JUL-1993.
PD
XX 17-JAN-1992; 92WO-US00427.
PF
XX 17-JAN-1992; 92WO-US00427.
PR
XX 17-JAN-1992; 92WO-US00427.
PA
XX (ABBO) ABBOTT LAB.
PI
XX Donadio S, Katz L, McAlpine JB;
XX WPI; 1993-242804/30.
DR
XX N-PSDB; AAQ46806.
N-PSDB; AAQ46806.

PT Biosynthesis of specific polyketide analogues esp. erythromycin
PT cpds. - by introducing altered biosynthetic gene-contg. DNA into
PT microorganisms
PS Disclosure; Fig 2; 133p; English.

XX The sequences given in AAR44430-32 are encoded by the eryA fragment of
CC the Saccharopolyspora erythraea genome. These polypeptides are
CC involved in the biosynthesis of the polyketide segment of erythromycin.
CC eryA is organized in modules and each module takes care of one
CC condensation step. The precise succession of elongation steps is
CC dictated by the genetic order of the modules. The DNA encoding
CC these polypeptides may be specifically altered such that novel
CC polyketide molecules of desired structure are produced. Three types
CC of alteration may be produced; those inactivating a single function in
CC a module which does not arrest acyl chain growth; those inactivating a
CC single function in a module which does affect chain growth; and those
CC affecting an entire module. The mutations may be introduced by gene
CC replacement.

SQ Sequence 3567 AA;

Query Match 55.4%; Score 884.5; DB 14; Length 3567;

Best Local Similarity 57.3%; Pred. No. 1.2e-77; Matches 177; Conservative 41; Mismatches 88; Indels 3; Gaps 2;

QY 1 VFVPPGCGGAGWGVRLASSPVFARRMQACEBALAPWVDSVVDILRRDAGDAWERA 60
DB 560 VFVPPGCGGAGWGVRLASSPVFARRMQACEBALAPWVDSVVDILRRDAGDAWERA 60
QY 61 DVVQPVLFSSVMSIALMRSYGIPEPAVLGHSGEIAAAHVCGALSLKDAKTVALLRSRA 120
DB 620 DVVQPVLFSSVMSIALMRSYGIPEPAVLGHSGEIAAAHVCGALSLKDAKTVALLRSRA 120
QY 121 LAAVRGRGMAVSLPDAOEVEQLIGERMAGRLVAAVNGPRSTVSGDAEAVDEVLYACA 180
DB 680 LAAVRGRGMAVSLPDAOEVEQLIGERMAGRLVAAVNGPRSTVSGDAEAVDEVLYACA 180
QY 181 GTGVARRIPVDYASHCPHVPOLREBELLELGDISPQSGVPPFSTVEGTWLDTTTLDA 240
DB 739 GTGVARRIPVDYASHCPHVPOLREBELLELGDISPQSGVPPFSTVEGTWLDTTTLDA 240
QY 241 YWYRNLIHQVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 300
DB 799 YWYRNLIHQVRFSDAVQALADGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 300
QY 299 RGDNDTRRF 307

Db 859 RDGGLSAP 867

RESULT 10

AAV92707
ID AAV92707 standard; Protein; 4150 AA.

AAV92707,

29-AUG-2000 (first entry)

S. antibiotic 8,8a-deoxyoleandolide synthase I encoded by oleAI.

oleandomycin, oleandolide; polyketide synthase; oleAI; oleAII; oleAIII;
PKS; type I; 8,8a-deoxyoleandolide synthase; modular; ketosynthase;
acyl-transferase; acyl carrier protein; inactivated; polyketide;
macrolactone; antibiotic; motilide; erythromycin.

Streptococcus antibioticus.

Location/Qualifiers

9..1034
/label= Loading_module

9..428
/label= KS-Q domain
/note= "Inactivated ketosynthase domain"

562..931
/label= AT domain
/note= "acyl carrier protein domain"

934..1034
/label= ACP domain
/note= "acyl carrier protein domain"

1061..1478
/label= Extender_module_1

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

1580..1926
/label= ketosynthase domain"

PI Belach MC, Shah SK, McDaniel R, Tang L;

XX WPI; 2000-365602/31.

DR N-PSDB; AAA09469.

PT Recombinant DNA compound encoding oleandolide polyketide synthase for

PS loading module or any one of extender modules

XX Disclosure; Page 27-28; 86pp; English.

CC The oleandolide polyketide synthase (PKS), also known as

CC 8,8a-deoxyoleandolide synthase, is encoded by three open reading frames

CC (ORF), designated oleAI, oleAII and oleAIII. The PKS is a type I

CC "modular" enzyme, where each ORF encodes 2 extender modules and

CC the first ORF also encodes the loading module. Each module is composed

CC of at least a ketosynthase (KS), acyl-transferase (AT) and an

CC acyl carrier protein (ACP) domain. The oleandolide PKS loading module

CC contains an inactivated KS, called KS-Q, where Q is the abbreviation for

CC glutamine, present instead of the active site cysteine required for

CC activity. The large multifunctional PKS enzymes catalyze the biosynthesis

CC of polyketide macrolactones through multistep pathways involving

CC decarboxylative condensations between acylthioesters followed by cycles

CC of varying beta-carbon processing activities. The macrolide product of

CC the PKS, 8,8a-deoxyoleandolide, is further modified by epoxidation and

CC glycosylation to yield oleandomycin, an antibacterial polyketide. The

CC invention concerns an isolated recombinant DNA compound, comprising a

CC coding sequence for a domain of loading module or any one of extender

CC modules 1-4 or 1-6, including an oleandolide PKS operably linked to a

CC promoter. Also discussed are recombinant oleandolide PKS in which the

CC module 1 KS domain is inactivated by deletion or other mutation. In

CC particular, the inactivation is mediated by a change in the KS domain

CC that renders it incapable of binding substrate (the KS-Q mutation). The

CC oleandolide PKS is useful for synthesizing polyketides, which are useful

CC as antibiotics and molluscs. Heterologous expression of oleandolide PKS

CC in host cells such as Streptomyces coelicolor and S. lividans is also

CC made possible. Unmodified oleandolide compounds can be provided to

CC cultures of Saccharopolyspora erythraea and converted to the

CC corresponding derivatives of erythromycin A-D.

XX Sequence 4150 AA;

SQ

Query Match 53.9%; Score 861.5; DB 21; Length 4150;

Best local Similarity 56.7%; Pred. No. 2, 8e-75;

Matches 174; Conservative 35; Mismatches 95; Indels 3; Gaps 2;

QY 1 VFVFGQSGQWAGVRLTASSVFPARRMOACBBLAPWVMSVVDILRRDAGDAWERA 60

Db 3076 VFVFGQSGQWAGVRLTASSVFPARRMOACBBLAPWVMSVVDILRRDAGDAWERA 60

QY 61 DVVQPVLFVSVVSLALMRSYGEEDPAVIGHSGGELAAHVACALSLKDAKTVLRSRA 120

Db 3134 DVVQPVLFVSVVSLALMRSYGEEDPAVIGHSGGELAAHVACALSLKDAKTVLRSRA 120

QY 121 LAAVRGRGMAAPLPAQEVOLIGRMWGRLLVVAVNGPRSPAVSGDAVAVBEVLA 180

Db 3194 LAAVRGRGMAAPLPAQEVOLIGRMWGRLLVVAVNGPRSPAVSGDAVAVBEVLA 180

QY 3253 REGVRRARVVDYASHSAGQMDQLRDLLEALDVTQDSVPFSTVTADMTTLADAG 3312

Db 181 GGVGRRARVVDYASHSAGQMDQLRDLLEALDVTQDSVPFSTVTADMTTLADAG 3312

QY 241 YWYRNLIHOPVRESDAVQALADGHRVYVVSPTLVPALEDTEDTAEVTAIGSLRG 300

Db 3313 YWYRNLIHOPVRESDAVQALADGHRVYVVSPTLVPALEDTEDTAEVTAIGSLRG 300

QY 301 DNDTRRF 307

Db 3373 EGGLDRF 3379

RESULT 11

ID	AA92709	standard; Protein; 3519 AA.
XX	AA92709	
AC	AA92709;	
DT	29-AUG-2000	(first entry)
XX		
DE	S. antibioticus 8,8a-deoxyoleandolide synthase III encoded by oleAIII.	
XX		
KW	oleandromycin; oleandrolide; polyketide synthase; oleA1; oleAII; oleAIII;	
KW	PKS; type I; 8,8a-deoxyoleandolide synthase; modular; ketosynthase;	
KW	acyl-transferase; acyl carrier protein; inactivated; polyketide;	
KW	macrolactone; antibiotic; motilide; erythromycin.	
XX		
OS	Streptococcus antibioticus.	
XX		
FH	Key	Location/Qualifiers
FT	Region	33..1564
FT		/label= Extender_module_5
FT	Domain	33..466
FT		/label= KS5_domain
FT		/note= "ketosynthase domain"
FT	Domain	569..915
FT		/label= AT5_domain
FT		/note= "acyl transferase domain"
FT	Domain	1199..1383
FT		/label= KR5_domain
FT		/note= "ketoreductase domain"
FT	Domain	1478..1564
FT		/label= ACP5_domain
FT		/note= "acyl carrier protein domain"
FT	Region	1686..3218
FT		/label= Extender_module_6
FT	Domain	1686..2115
FT		/label= KS6_domain
FT		/note= "ketosynthase domain"
FT	Domain	2220..2566
FT		/label= AT6_domain
FT		/note= "acyl transferase domain"
FT	Domain	2855..3039
FT		/label= KR6_domain
FT		/note= "ketoreductase domain"
FT	Domain	3132..3218
FT		/label= ACP6_domain
FT		/note= "acyl carrier protein domain"
FT	Domain	3290..3519
FT		/label= TR_domain
FT		/note= "thiol esterase domain"
XX		
PN	WO200026349-A2.	
XX		
PD	11-MAY-2000.	
XX		
PE	22-OCT-1999;	99WO-US24478.
XX		
XX	29-OCT-1998;	98US-0106100.
PR	16-FEB-1999;	99US-0120254.
XX		
PA	(KOSA-) KOSAN BIOSCIENCES INC.	
XX		
PI	Betlach MC, Shah SK, McDaniel R, Tang L;	
XX		
DR	WPI; 2000-365602/31.	
XX	N-PSDB; AAA09469.	
FT	Recombinant DNA compound encoding oleandrolide polyketide synthase for	
FT	synthesizing polyketides comprising a coding sequence for a domain of a	
FT	loading module or any one of extender modules	
XX		
PS	Disclosure; Page 29; 86pp; English.	
XX		
CC	The oleandrolide polyketide synthase (PKS), also known as	
CC	8,8a-deoxyoleandolide synthase, is encoded by three open reading frames	

(ORF), designated oleA1, oleA2 and oleA3. The PKS is a type I "modular" enzyme, where each ORF encodes 2 extender modules and the first ORF also encodes the loading module. Each module is composed of at least a ketosynthase (KS), acyl-transferase (AT) and an acyl carrier protein (ACP) domain. The oleandrolide PKS loading module contains an inactivated KS, called KS-Q, where Q is the abbreviation for glutamine, present instead of the active site cysteine required for activity. The large multifunctional PKS enzymes catalyze the biosynthesis of polyketide macroketones through multistep pathways involving decarboxylative condensations between acylthioesters followed by cycles of varying beta-carbon processing activities. The macroide product of the PKS, 8,8a-deoxyoleandrolide, is further modified by epoxidation and glycosylation to yield oleandomycin, an antibacterial polyketide. The invention concerns an isolated recombinant DNA compound comprising a coding sequence for a domain of loading module or any one of extender modules 1-4 or 1-6, including an oleandrolide PKS operably linked to a promoter. Also discussed are recombinant oleandrolide PKS in which the module 1 KS domain is inactivated by deletion or other mutation. In particular, the inactivation is mediated by a change in the KS domain that renders it incapable of binding substrate (the KS1-Q mutation), rendered by mutation in the codon for the active site cysteine. The oleandrolide PKS is useful for synthesizing polyketides, which are useful as antibiotics and mollusids. Heterologous expression of oleandrolide PKS in host cells such as *Streptomyces coelicolor* and *S. lividans* is also made possible. Unmodified oleandrolide compounds can be provided to cultures of *Saccharopolyspora erythraea* and converted to the corresponding derivatives of erythromycins A-D.

Query Match	53.7%;	Score 858;	DB 21;	Length 3519;
Best Local Similarity	57.0%;	Pred. No. 4.9e-75;		
Matches 175;	Conservative 34;	Mismatches 94;	Indels 4;	Gaps 3;

QY	I	VFVPRGGGGAAGGVLLASSPFAARMQACEALAPWMDSVYDLR-RDAGAVWER	59
Db	2221	VFFVPPGGGSQWVGAAGGLDACPFAFAEABCAALVDVTCMSLVEVLOGRDA--TVLGR	227/8
QY	60	ADVVPQVLFVYVMSLAAIMRSYGLIEPDAVLGHSQGEIILAAHYCGALSLDKDAKTVALRSR	119
Db	2279	VDVVPQALMAVWVSLARTWRYGVGEPAVAVGHSQGEIILAAVAGGLSLADBARVVLRSR	233/8
QY	120	ALLAVRGGGGAASVPLPAQVEIQLIGERWAGRLLWVAANGRSTAVSGDAARVDEVLYC	179
Db	2339	ALARTAGGGGVAWSLSPAGRVRTML-DYIGGRSVVAANGSSSTVSSGDDVQULDELLAGC	239
QY	180	AGTVGARARILPVDYASHCHPHVQPIREBELLLELLGISIPQSVPPFSFIVEGTWLDPTTLDA	239
Db	2398	EREGRGARVRVPDYASHQMDQLRELLLELLAALITPDQSSVPFFSITVALDMLDTTALDA	245/8
QY	240	AYWYRNLIHQPRFSDAYLADDDGHRVVEVSPHTLVPALEDTBEDPAVDTVALGISLR	299
Db	2458	GYTFPTNLRETYRFQDAVEGLVAQGMGAFAVESPPHVLVPGLQTLDDALDQNAAVLGISLR	251/1
QY	300	GDNDTRR 306	
Db	2518	DEGGIDR 2524	

RESULT 12
AAV39297
ID AAV39297 standard; Protein; 2595 AA.
XX
XX AAV39297;
AC
AC
XX
DT 01-DEC-1999 (first entry)
XX
DE Spma a polyketide synthase.
XX
XX Spinosyn biosynthetic enzyme; open reading frame; ORF; insecticidal
KW microides; arachnid; nematode; insect; polyketide; polyketide synthase;
KW PKS; extender module; initiator module; acyl transferase domain; AT;
KW acyl carrier protein; ACP; beta-ketosynthase domain; KS; KR;

KM		dehydratase domain; DH; enoyl reductase domain; ER; beta-ketoreductase;
KW		insecticide.
XX		
OS		Saccharopolyspora spinosa.
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference	1944
FT	/note=	"Encoded by CAG"
FT	/label=	KSI
FT	/note=	"Beta-ketosynthase domain: part of the initiator module"
FT	Domain	528..853
FT	/label=	Art1
FT	/note=	"Acyl transferase domain: part of the initiator module"
FT	Domain	895..977
FT	/label=	ACPI
FT	/note=	"Acyl carrier protein domain: part of the initiator module"
FT	Domain	998..1413
FT	/label=	KSI
FT	/note=	"Beta-ketosynthase domain: part of extender module 1"
FT	Domain	1525..1858
FT	/label=	AT1
FT	/note=	"Acyl transferase domain: part of extender module 1"
FT	Domain	2158..2337
FT	/label=	KR1
FT	/note=	"Beta-ketoreductase domain: part of extender module 1"
FT	Domain	2432..2513
FT	/label=	ACPI
FT	/note=	"Acyl carrier protein domain: part of extender module 1"
XX		
PN	WO9946387-A1.	
XX		
PD	16-SEP-1999.	
XX		
PF	16-FEB-1999;	99MO-US03212.
XX		
PR	09-MAR-1998;	98US-0036987.
PA	(DOMC) DOW AGROSCIENCES LLC.	
XX		
PI	Baltz RH, Broughton MC, Crawford KP, Madhuri K, Merlo DJ,	
PI	Treadway PJ, Turner JR, Waldron CJ;	
DR	WPI; 1999-551414/46.	
N-PSDB	AZ21501.	
XX		
PT	New spinosyn biosynthetic genes from Saccharopolyspora spinosa, useful	
PT	for production of insecticidal spinosyn compounds -	
PS	Claim 1; Page 87-96; 190pp; English.	
XX		
CC	This is the amino acid sequence of the product of the spnA gene. The	
CC	protein is involved in spinosyn biosynthesis. The spnA gene is one of 233	
CC	genes and open reading frames contained in an 80kb DNA sequence	
CC	AZ21501. Spinosyns are insecticidal microclides which are useful for the	
CC	control of arachnids, nematodes and insects. Biosynthesis of spinosyns	
CC	occurs via stepwise condensation and modification of carboxylic acid	
CC	precursors generating a linear polyketide which is modified further. These	
CC	DNA sequence contains a central region of approximately 55kb which has	
CC	homology to the DNA encoding the polyketide synthases (PKS) of known	
CC	microclide producers. The spinosyn PKS DNA region consists of 5 ORFs with-	
CC	stop codons at the end of acyl carrier protein (ACP) domains. Together	
CC	the PKS polypeptides (AAVY9297-Y39301) of which this sequence is one,	
CC	form a complex consisting of an initiator module, spnA, and several	
CC	extender modules spnB-spnE. Each extender module adds a specific acetyl	
CC	Co-A precursor to a growing polyketide chain, and modifies the beta-keto	

CC group in a specific manner. A module in a PKS polypeptide consists of several domains with specific functions. The initiator module has an acyl transferase (AT) domain, and an acyl carrier protein (ACP) domain. The extender modules have the same domains plus a beta-keto synthase (KS) domain and optionally a beta-ketoreductase domain, a dehydratase (DH) domain, and an enoyl reductase (ER) domain. The last extender module terminates with a thioester domain. The products of the genes present in the upstream region the PKS genes have been assigned names spnF-spnS AA19302-Y39315 and are responsible for different modifications in spinosyn biosynthesis. There are also two ORFs ORP15 and ORP16 present immediately upstream of spns, producing polypeptides AA19316-Y39317, and two ORFs ORP1 and ORP2 present downstream of the PKS region producing polypeptides AA19318-Y39319. The genes are useful to improve yields of spinosyn, and for creating new spinosyns e.g. by mutagenesis, or interruption of steps in spinosyn biosynthesis. The modified spinosyns may be a new insect control agent or serve as substrates for further chemical modification and the creation of new semi-synthetic spinosyns. The genes are also useful to isolate similar sequences from S. spinosa or other species by hybridization.

SQ Sequence 2595 AA;

Query Match 53.3%; Score 852; DB 20; Length 2595;
Best Local Similarity 55.2%; Pred. NO. 1.2e-74;
Matches 169; Conservative 47; Mismatches 86; Indels 2; Gaps 2

QY 1 VFVFPGGCGMAGNGVRLASPPVFARRMOACEALAPWVDMSVVDILRRDAGDAWERRA 60
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 527 VFFVPFGGSGMGIGRQLLETSEVFASMSGDCAALAPHLDMSLDLVRNNAAGAHLHDH 586
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 61 DVVQPVLFSSVWSLALMRSTGYEPDAVLGHSGCEIAAAHYGCALSCLKDAAKTVALRSRA 120
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 587 DVVQPAFAFMVSLSLELMRSMGVRFVVVHGSGEIAAACACAGALSVDADARVVAVR SRL 646
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 121 LAAVRGGRGNASVPLPAOEVEOLIGERMAGRLLWAAANGPRSTVSGDAEAVDEVLAYCA 180
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 647 LTLAASGANASLOHPAEVERÖLL-LPMRDPIGVAGVGPSSTLVSGREMAELMECA 705
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 181 GTGVARRIRIVDVASHCPHVQPLREELLLELGDISPOPSGVPFESTEGTWLDITTTLPAA 240
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 706 DRELRRRIIVEYASHSPHLEVVRDELGLLAPVPERPTGSPIYSTTTTGDDLD-RPMAD 764
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 241 YWRNLHQVRFSDAVQALADDGHRFVEVSPHEFLVPAIEDTTEDTAEDVTAGSLRG 300
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 765 YWRNLHQVPLFEAAVEALLRGYDAFIETISPHVLTRANIQTAVRAGREVVALGTIARG 824
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 301 DNDDR 306
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 825 EGGMRQ 830

RESULT 13
AAB70965
ID ID AAB70965 standard; Protein: 2595 AA.
XX AC AAB70965;
DT 28-AUG-2001 (first entry)
XX XX
DE S. epinosia protein fragment encoded by ORF18, SEQ ID 41.
XX XX
XX Porosamine; trimethylchamnose; polyketide synthase; biosynthesis;
KM spinosyn; polyketide aglycone; transgenic plant; insect resistance;
KW macroide; insecticidal; polyketide synthase.
OS Saccharopolyspora spinosa.
XX DE19957268-A1.
XX PN 08-MAR-2001.
XX PF 29-NOV-1999; 99DE-1057268.

PR 27-AUG-1999; 99DE-1040596.
 XX (FARB) BAYER AG.
 XX Eberz G, Moehrle V, Froede R, Velten R, Salas JA,
 XX WPI; 2001-267102/28.
 DR N-PSDB; AAF68335.
 PT New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
 PT recombinant production of insecticidal spinosyns and their derivatives
 PT
 PS Claim 57; Page 179-189; 354pp; German.
 XX
 CC This invention describes a novel method nucleic acid (I) and its encoded
 CC polypeptide (II) containing at least one region that encodes an enzymatic
 CC activity involved in biosynthesis of spinosyns. (I) are used (1) to
 CC identify, inactivate or modulate genes involved in the biosynthesis of
 CC (II); (ii) to generate a library of polypeptide synthetases; (iii) for
 CC adding forosamine or trimethylrhamsone to a spinosyn or polypeptide
 CC aglycone; and (iv) for recombinant production of the corresponding
 CC enzymes, which are used for production of (II), their precursors or
 CC derivatives, including production of transgenic plants that express (II)
 CC and thus have increased resistance to insects. (I) are also useful as
 CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
 CC macrocides with insecticidal, but not antibacterial, activity, and can
 CC also be used to raise specific antibodies, useful for identifying
 CC expression clones in a gene bank. Cells transformed with (I) may produce
 CC (II) at significantly increased levels or produce new derivatives of
 CC (II). This sequence represents an S. spinosa polypeptide synthetase.
 XX
 SQ Sequence 2595 AA;
 Query Match 53.3%; Score 852; DB 22; Length 2595;
 Best Local Similarity 55.2%; Pred. No. 1.2e-74;
 Matches 169; Conservative 47; Mismatches 88; Indels 2; Gaps 2;
 QY 1 VVFPFGGGGGMAGVGLLASSPFARRMQACERALLPMDWSVVDLIRDAGDAWERA 60
 DB 527 VFVFPFGGGGSMIGRQLLETSEVFAASMSDCADALPHLDWSLIDVLRNAAGAHLDHD 586
 QY 61 DVVQPVLFVWVSLAALMRSGIEPDAVLGHSGEIAAAHVCALSLDKAKTVALLSRA 120
 DB 587 DVVQPALFALMVSLAELMRSGVPRVAVVGHSGEIAAACVAGLSRDAARVAVASRL 646
 QY 121 LAAVRGSGVAVPLPAQVEVQLIGERWAGRLWVAANVGPSTAVSGDAEAVDEVLAYCA 180
 DB 647 LTLAAGSGAMASLQHPABEVQIL-LPMRDRIGVAGVNGPSSSTVSGDREMAELLAECA 705
 QY 181 GTGVRARIPVDVASHCPHQPLREELLGLDISQPSGVPRFSTVEGTMLOTTIDAA 240
 DB 706 DRELRMRIPVEVASHSPHLEVDLGLLAPPEPTGSPITSTTGDLID-RPMADAD 764
 QY 241 YWYRNLIHQVRFSDAVOALADGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 300
 DB 765 YWYRNLIHQVRFSDAVOALADGHRVFEVSPHPTLVPAIEDTETDAEDVTAIGSLRG 824
 QY 301 DNDTRR 306
 DB 825 EGGWRQ 830

RESULT 14

AAV39299
 ID AAV39299 standard; Protein; 3170 AA.
 XX
 AC AAV39299;
 XX
 DT 01-DEC-1999 (first entry)
 XX
 DE SpnC a polypeptide synthetase.
 XX

KW Spinosyn biosynthetic enzyme; open reading frame; ORF; insecticidal
 KW microtides; arachnid; nematode; insect; polypeptide; polypeptide synthetase;
 KW PKS; extender module; initiator module; acyl transferase domain; AT;
 KW acyl carrier protein; ACP; beta-ketosynthase domain; KS; KR;
 KW dehydratase domain; DH; enoyl reductase domain; ER; beta-ketoreductase;
 KW insecticide.
 XX
 XX Saccharopolyspora spinosa.
 OS
 FH Key Location/Qualifiers
 FT Domain 1..423
 FT /label= KS3
 FT /note= "Beta-ketosynthase domain: part of extender
 FT module 3"
 FT Domain 531..850
 FT /label= AT3
 FT /note= "Acyl transferase domain: part of extender
 FT module 3"
 FT Domain 1159..1337
 FT /label= KR3
 FT /note= "Beta-ketoreductase domain: part of extender
 FT module 3"
 FT Domain 1425..1506
 FT /label= ACP3
 FT /note= "Acyl carrier protein domain: part of extender
 FT module 3"
 FT Domain 1529..1952
 FT /label= KS4
 FT /note= "Beta-ketosynthase domain: part of extender
 FT module 4"
 FT Domain 2066..2396
 FT /label= AT4
 FT /note= "Acyl transferase domain: part of extender
 FT module 4"
 FT Domain 2700..2880
 FT /label= KR4
 FT /note= "Beta-ketoreductase domain: part of extender
 FT module 4"
 FT Domain 2972..3053
 FT /label= ACP4
 FT /note= "Acyl carrier protein domain: part of extender
 FT module 4"
 FT Domain MO9946387-AL.
 PD 16-SEP-1999.
 XX
 PF 16-FEB-1999; 99WO-US03212.
 XX
 PR 09-MAR-1998; 98US-0036987.
 XX
 PA (DOWC) DOW AGROSCIENCES LLC.
 XX
 PI Baltz RH, Broughton MC, Crawford KP, Madduri K, Merlo DJ,
 PI Treadway PJ, Turner JR, Waldron C;
 XX
 DR WPI; 1999-551414/46.
 DR N-PSDB; AA221501.
 XX
 XX New spinosyn biosynthetic genes from Saccharopolyspora spinosa, useful
 PT for production of insecticidal spinosyn compounds
 PT
 PS Claim 1; Page 103-113; 190pp; English.
 XX
 CC This is the amino acid sequence of the product of the spnC gene. The
 CC protein is involved in spinosyn biosynthesis. The spnC gene is one of 23
 CC genes and open reading frames contained in an 80kb DNA sequence
 CC AA221501. Spinosyns are insecticidal microtides which are useful for the
 CC control of arachnids, nematodes and insects. Biosynthesis of spinosyns
 CC occurs via stepwise condensation and modification of carboxylic acid
 CC precursors generating a linear polypeptide which is modified further. The
 CC DNA sequence contains a central region of approximately 55kb which has
 CC homology to the DNA encoding the polypeptide synthetases (PKS) of known

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: June 17, 2003, 12:53:02 ; Search time 183.577 Seconds
(without alignments)
4458,180 Million cell updates/sec

Title: US-09-914-286-3
Perfect score: 20829
Sequence: 1 VQRMDSGEPRPAPAGEVLGV.....FTSATBAEIRPKFIDNDIGLS 3972

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	20826	100.0	3972	2	Q9S0R8 streptomyc
2	17634.5	84.7	3626	2	Q9F779 streptomyc
3	12972	62.3	5532	2	Q9S0R4 streptomyc
4	11330.5	54.4	6239	2	Q9S0R7 streptomyc
5	8148	39.1	4613	2	Q9ZG15 streptomyc
6	7891	37.9	6048	2	Q93H87 streptomyc
7	7705.5	37.0	4472	2	Q93954 streptomyc
8	7516.5	36.1	6146	2	Q93H05 streptomyc
9	7356	35.3	5588	2	Q93H01 streptomyc
10	7281.5	35.0	4340	2	Q9ALM2 streptomyc
11	7244	34.8	4340	2	Q93H01 streptomyc
12	7221.5	34.7	6797	2	Q9X993 streptomyc
13	7217	34.6	4840	2	Q93H14 streptomyc
14	7190.5	34.5	4881	2	Q9S0R3 streptomyc
15	7168	34.4	10917	2	Q93NM6 streptomyc
16	7098	34.1	5069	2	Q52789 amycolatops

17	7094	34.1	5060	2	Q52545 amycolatops
18	7089.5	34.0	11036	2	Q914W3 streptomyc
19	7075	34.0	4498	2	Q93H85 streptomyc
20	7049.5	33.8	4290	2	Q9WXC0 micromonsp
21	7041.5	33.8	3816	2	Q9KTV3 streptomyc
22	7017	33.7	3739	2	Q9ZG14 streptomyc
23	6986	33.5	3939	2	Q93H03 streptomyc
24	6952.5	33.5	3576	16	Q9EX53 streptomyc
25	6943.5	33.3	4685	2	Q93H02 streptomyc
26	6939.5	33.3	5435	2	Q914X2 streptomyc
27	6912.5	33.2	4150	2	Q9KTV4 streptomyc
28	6828	32.8	6155	2	Q93H84 streptomyc
29	6813.5	32.7	4151	16	Q53490 mycobacteri
30	6804.5	32.7	3170	2	Q9ALM4 streptomyc
31	6778.5	32.5	3524	2	Q93H86 streptomyc
32	6756	32.4	10223	2	Q54296 streptomyc
33	6753	32.4	9507	2	Q9EWA1 streptomyc
34	6728	32.3	8817	2	Q53840 streptomyc
35	6716.5	32.2	7525	2	Q9K1E0 streptomyc
36	6652	31.9	9477	2	Q914X3 streptomyc
37	6642.5	31.9	8563	2	Q54297 streptomyc
38	6624.5	31.8	3970	2	Q93H18 streptomyc
39	6585.5	31.7	9510	2	Q93HX9 streptomyc
40	6579.5	31.6	3654	2	Q30766 streptomyc
41	6555	31.5	3729	2	Q33956 streptomyc
42	6469.5	31.1	5644	2	Q93NX8 streptomyc
43	6447	31.0	4928	2	Q9ALM3 streptomyc
44	6416	30.8	7576	2	Q9ZG44 streptomyc
45	6302.5	30.3	3546	2	Q9P830 micromonsp

ALIGNMENTS

RESULT 1

Q9S0R8 PRELIMINARY; PRT; 3972 AA.
AC Q9S0R8; 01-MAY-2000 (TRENBLER. 13, Created)
DT 01-MAY-2000 (TRENBLER. 13, Last sequence update)
DT 01-MAR-2002 (TRENBLER. 20, Last annotation update)
DE Type I polypeptide synthase AVES 1.
GN AVEA1.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9380548; Pubmed=10449723;
RA Ikeda H., Nonomura T., Usami M., Ohta T., Omura S.;
RT "Organization of the biosynthetic gene cluster for the polypeptide
antihelminthic macrocyclic avermectin in Streptomyces avermitilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514(1999).
DR EMBL; AB032367; BAA84474.1; -
DR InterPro; IPR00127; Ac transferase.
DR InterPro; IPR000794; Keloacyl-synt.
DR InterPro; IPR003880; Prantine-attach.
DR InterPro; IPR000834; Zn-carboxypept.
DR Pfam; PF00698; Acyl_transf. 3.
DR Pfam; PF00109; ketoacyl-synt. 2.
DR Pfam; PF02801; ketoacyl-synt. C; 2.
DR Pfam; PF00550; pp-binding. 3.
DR PROSITE; PSS0075; ACP DOMAIN. 3.
DR PROSITE; PSS0066; B_KETOACYL_SYNTHASE. 2.
DR PROSITE; PSS00133; CARBOXYPEPT_ZN_2; UNKNOWN_2.
DR PROSITE; PSS00012; PHOSPHOPANTETHEINE; 2.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 3972 AA; 416852 MW; 2A2936958032B1C3 CRC64;

Query Match 100.0%; Score 20826; DB 2; Length 3972;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3971; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VORMDGEPRPAAGEVLGVADADGVFVFPQGPOMPGMRELLDASDVFRSEVAC 60
Db 1 MORMDGEPRPAAGEVLGVADADGVFVFPQGPOMPGMRELLDASDVFRSEVAC 60
QY 61 EAAFPYVDWSEVOYLKDS PDAPGLDRVDVVOPTLFAVMI SIALMWSQVPCAVLGH 120
Db 61 EAAFPYVDWSEVOYLKDS PDAPGLDRVDVVOPTLFAVMI SIALMWSQVPCAVLGH 120
QY 121 LGETIAAHVSGSLADAAVVTLMSSOQTTLAGALVSVAATPDELLPRLAPTEPNP 180
Db 121 LGETIAAHVSGSLADAAVVTLMSSOQTTLAGALVSVAATPDELLPRLAPTEPNP 180
QY 181 ARLAAVAVNGPRSTVSGAREAVADLVADLTAQVRTMIPVDVAHSPMLYALHERVVS 240
Db 181 ARLAAVAVNGPRSTVSGAREAVADLVADLTAQVRTMIPVDVAHSPMLYALHERVVS 240
QY 241 GLLPITPPRSRI PFHSSVTGRLDRELDAAYVYRNMSTVRFEBARLLLOQGEKTFE 300
Db 241 GLLPITPPRSRI PFHSSVTGRLDRELDAAYVYRNMSTVRFEBARLLLOQGEKTFE 300
QY 301 MSPHVLTMGOELAPDLGDTTGTADYIMGTLRGOCTTLDFLTLAQLAGHETSAT 360
Db 301 MSPHVLTMGOELAPDLGDTTGTADYIMGTLRGOCTTLDFLTLAQLAGHETSAT 360
QY 361 VLSARLTALSPTOOQSLLDLVRAHTMAVLNDGNERSTADGSPASFAHLGFDSVMGVE 420
Db 361 VLSARLTALSPTOOQSLLDLVRAHTMAVLNDGNERSTADGSPASFAHLGFDSVMGVE 420
QY 421 LRNLRSKATGLRLPYTLIFDHTTPAAVAAARLRTAALGHLDEDTAEPVDSPSGHGTAAD 480
Db 421 LRNLRSKATGLRLPYTLIFDHTTPAAVAAARLRTAALGHLDEDTAEPVDSPSGHGTAAD 480
QY 481 DPILAIGACRPGCVSRPKDLMELAA SGDAIGFPDRGMPTROBAROAPTOPTGTYR 540
Db 481 DPILAIGACRPGCVSRPKDLMELAA SGDAIGFPDRGMPTROBAROAPTOPTGTYR 540
QY 541 QGGGFLHDAHFDAGFFGISPREALAMPQORLLLETWEAFERAGIDPLSVRSRTVF 600
Db 541 QGGGFLHDAHFDAGFFGISPREALAMPQORLLLETWEAFERAGIDPLSVRSRTVF 600
QY 601 AGALSPDYGPMDTASSEGAADVEGHILGTGTSVLSGRILAYSGLEGPALTVDYDGSAS 660
Db 601 AGALSPDYGPMDTASSEGAADVEGHILGTGTSVLSGRILAYSGLEGPALTVDYDGSAS 660
QY 661 LVTLHLCOSLRSGECTLALAGVSVKSTLGMTEFSGORGLSVDRGKKAISAADGTGW 720
Db 661 LVTLHLCOSLRSGECTLALAGVSVKSTLGMTEFSGORGLSVDRGKKAISAADGTGW 720
QY 721 GEGVGMILVERLSDAVRLGHRVLAVVRGSAVNODGASNGLTAPNCPAORVITROLANAG 780
Db 721 GEGVGMILVERLSDAVRLGHRVLAVVRGSAVNODGASNGLTAPNCPAORVITROLANAG 780
QY 781 LSVADVVDVEGHGTGTLGDPLEAQAALLATYGORAGORPLWLGSLKSNIGHTMAAGVGG 840
Db 781 LSVADVVDVEGHGTGTLGDPLEAQAALLATYGORAGORPLWLGSLKSNIGHTMAAGVGG 840
QY 841 VIKVMMLRBEGLVRLTHVDKPSQVMSAGAVLLTFAVWMPDAGRLRRAVSSFGI 900
Db 841 VIKVMMLRBEGLVRLTHVDKPSQVMSAGAVLLTFAVWMPDAGRLRRAVSSFGI 900
QY 901 GGTNAHVILBEAPAAAGCVAGGVLEGAFLAISVAESVAAPVAVSAVPVAVPVP 960
Db 901 GGTNAHVILBEAPAAAGCVAGGVLEGAFLAISVAESVAAPVAVSAVPVAVPVP 960
QY 961 VPVVSARSSEAGLQAQEAALRQYVAVRPDVSILADVAGLACGRAVLEHRAVLADEBEL 1020
Db 961 VPVVSARSSEAGLQAQEAALRQYVAVRPDVSILADVAGLACGRAVLEHRAVLADEBEL 1020
QY 1021 VOGIGALAAAGBDRRVTYTTGHA PGDGRGVVFPFGOGGONAGMVRLLASSPVARMOA 1080
Db 1021 VOGIGALAAAGBDRRVTYTTGHA PGDGRGVVFPFGOGGONAGMVRLLASSPVARMOA 1080

QY 1081 CEEALAPWDMVSVDILRRDAGDAWBERADVQVFLFSVMVSLAALMRSYGEIPDAVLGH 1140
Db 1081 CEEALAPWDMVSVDILRRDAGDAWBERADVQVFLFSVMVSLAALMRSYGEIPDAVLGH 1140
QY 1141 SGEIILAAHVCGALSLKDAKTTVALRSRALAANVRGCMASVPLPAOEVEOLIGERVAGR 1200
Db 1141 SGEIILAAHVCGALSLKDAKTTVALRSRALAANVRGCMASVPLPAOEVEOLIGERVAGR 1200
QY 1201 LMVAANVGPSTANSGAEAVDETLACAGGVARARIPVDYASHCHVOPRLRELELL 1260
Db 1201 LMVAANVGPSTANSGAEAVDETLACAGGVARARIPVDYASHCHVOPRLRELELL 1260
QY 1261 GDISPQSGVPFESTVEGTLDTTTLDAAYWYRNLHOPVFPSPAVQALADGHRVFEVS 1320
Db 1261 GDISPQSGVPFESTVEGTLDTTTLDAAYWYRNLHOPVFPSPAVQALADGHRVFEVS 1320
QY 1321 PHPTLVPAIEDTTEDETVTAIGSLRGNDTRRFLTALAHTTGTIGPTTWHHHTH 1380
Db 1321 PHPTLVPAIEDTTEDETVTAIGSLRGNDTRRFLTALAHTTGTIGPTTWHHHTH 1380
QY 1381 HHTHPHPTLIDLPTYRPOHOMWLESOPGAGSGGAGSGAGSGRAGTAGTAEVES 1440
Db 1381 HHTHPHPTLIDLPTYRPOHOMWLESOPGAGSGGAGSGAGSGRAGTAGTAEVES 1440
QY 1441 RFMDAVARODLETVAATLAVPSAGLDTVPVPAISAMRHODARINTWTYOETWKEPLTL 1500
Db 1441 RFMDAVARODLETVAATLAVPSAGLDTVPVPAISAMRHODARINTWTYOETWKEPLTL 1500
QY 1501 PTHQPHQWLIAPETOTHPHITNLINLHNGITPPLTLNHTHTNPOHLHTLHHT 1560
Db 1501 PTHQPHQWLIAPETOTHPHITNLINLHNGITPPLTLNHTHTNPOHLHTLHHT 1560
QY 1561 ROQAONHTGALIGLSLALDETPHHPHHTPTGTLNLTLOTHTQHPPTPLWAT 1620
Db 1561 ROQAONHTGALIGLSLALDETPHHPHHTPTGTLNLTLOTHTQHPPTPLWAT 1620
QY 1621 NATTTNPNDELTPHTQOTWGLARTTLLEHPTAGIIDLPTTPHTLOHLTOTLQPH 1680
Db 1621 NATTTNPNDELTPHTQOTWGLARTTLLEHPTAGIIDLPTTPHTLOHLTOTLQPH 1680
QY 1681 HOTQALRTGTHTRRLPTTLTPHTQPPRPHGTTLLIGTGALATHLTHLTHQPT 1740
Db 1681 HOTQALRTGTHTRRLPTTLTPHTQPPRPHGTTLLIGTGALATHLTHLTHQPT 1740
QY 1741 OHLLTSRGPHTPHQHLTOTLOQKGINHTITTCOTSNDOLOOLNTIPROHPLTTVI 1800
Db 1741 OHLLTSRGPHTPHQHLTOTLOQKGINHTITTCOTSNDOLOOLNTIPROHPLTTVI 1800
QY 1801 HTAGIILDDATLTNLTPTQLNNVLRKAHSAHLHQLTOHTPLTAFVLYSSAAATFGAPQ 1860
Db 1801 HTAGIILDDATLTNLTPTQLNNVLRKAHSAHLHQLTOHTPLTAFVLYSSAAATFGAPQ 1860
QY 1861 ANYAANAAYDALAHHRHTHLRPTSIAWGTCWONGIADSDKARAYLDRRGFPMSBELA 1920
Db 1861 ANYAANAAYDALAHHRHTHLRPTSIAWGTCWONGIADSDKARAYLDRRGFPMSBELA 1920
QY 1921 TAAVTOAILATERPVYVIAIDMSKIEHTSQTSDLVSAAREPAPVORPPPALHKTTLA 1980
Db 1921 TAAVTOAILATERPVYVIAIDMSKIEHTSQTSDLVSAAREPAPVORPPPALHKTTLA 1980
QY 1981 HOTSADORAALLLELVDRHVAALVRHADPKAIAIDQSFALGFPSTLAVEFRNLLIKATGL 2040
Db 1981 HOTSADORAALLLELVDRHVAALVRHADPKAIAIDQSFALGFPSTLAVEFRNLLIKATGL 2040
QY 2041 RLFPVSLFDPHTPAKLAHVHIONOLRGTAESAPSAAAVTAEASVTEPIAIVGMACFPGG 2100
Db 2041 RLFPVSLFDPHTPAKLAHVHIONOLRGTAESAPSAAAVTAEASVTEPIAIVGMACFPGG 2100
QY 2101 VTSADPFWDLISSEODAIIGFPPTDRGMDLDTLVPDPDHGTCTYTRNGGLLXDAGHFDAE 2160
Db 2101 VTSADPFWDLISSEODAIIGFPPTDRGMDLDTLVPDPDHGTCTYTRNGGLLXDAGHFDAE 2160
QY 2161 FFGISPREALAMPQORLLLETAMETIEHAGINPHTLHGPTGVFTGTNGODYALAVHNA 2220

```

Db      2161  FFGISREALAMPQOQRLLEETMETTEHGINPHTHGTPTGFTCTONQDVALRCHNA 2220
Qy      2221  GQSDGALGTGDSVSGRISYTFEGEPVAVSDTACSSLYVALHACOLARAGECSMA 2280
Db      2221  GQSDGALGTGDSVSGRISYTFEGEPVAVSDTACSSLYVALHACOLARAGECSMA 2280
Qy      2281  LAGGVTVMSSPGAFVEFSPQRGLADGCKAFSAADGTGMEGVMLVERLSDAHRNG 2340
Db      2281  LAGGVTVMSSPGAFVEFSPQRGLADGCKAFSAADGTGMEGVMLVERLSDAHRNG 2340
Qy      2341  HRVLAIVRGSADVNDGASNGLTAPNGPSQQRVTRQALANAGLADVDVANGSTTLLG 2400
Db      2341  HRVLAIVRGSADVNDGASNGLTAPNGPSQQRVTRQALANAGLADVDVANGSTTLLG 2400
Qy      2401  DPTBAQALMTYGGDRAGEBGLMGSVSNVGHQAAGVAVKYMVALRHGLPPTLH 2460
Db      2401  DPTBAQALMTYGGDRAGEBGLMGSVSNVGHQAAGVAVKYMVALRHGLPPTLH 2460
Qy      2461  VDERSPHVDMSAGAVQLLTETVPWPGEGRLRAGVSPGVSGTNAHVILEAPADVP 2520
Db      2461  VDERSPHVDMSAGAVQLLTETVPWPGEGRLRAGVSPGVSGTNAHVILEAPADVP 2520
Qy      2521  GPPAGEGDAGSDDEAAAGSPGWPMVLVSAKSQPALRAQALNHLTDHFGDLADVGT 2580
Db      2521  GPPAGEGDAGSDDEAAAGSPGWPMVLVSAKSQPALRAQALNHLTDHFGDLADVGT 2580
Qy      2581  LAHARAIVFDRATLIAADROTFLQALQALAAEPHRAVHSSANGCTEBAAGKTFIC 2640
Db      2581  LAHARAIVFDRATLIAADROTFLQALQALAAEPHRAVHSSANGCTEBAAGKTFIC 2640
Qy      2641  SGQGTQPGMAHGLYHTRPVPAALANDICTHLDPHLHPLPPLTQNDNEDAAALLOQ 2700
Db      2641  SGQGTQPGMAHGLYHTRPVPAALANDICTHLDPHLHPLPPLTQNDNEDAAALLOQ 2700
Qy      2701  TRYAOPALFARQVALHRLTDGYHITPHYVAGHSLGEITAAHLAAGILTLDTATTLTORA 2760
Db      2701  TRYAOPALFARQVALHRLTDGYHITPHYVAGHSLGEITAAHLAAGILTLDTATTLTORA 2760
Qy      2761  TLYQMTPRGTMTHLTHPHNTHNLTAHENDLAAINTPRLSVISGTPHVOHITTLCO 2820
Db      2761  TLYQMTPRGTMTHLTHPHNTHNLTAHENDLAAINTPRLSVISGTPHVOHITTLCO 2820
Qy      2821  QQGIKTPLPTNNAFHSPTNPILNQHQHTQTLTPHPTPLTANTPPDQLTPHXT 2880
Db      2821  QQGIKTPLPTNNAFHSPTNPILNQHQHTQTLTPHPTPLTANTPPDQLTPHXT 2880
Qy      2881  QOARNTVDYATTTQTLHQHGVTVYIELGRDNTLTTLTHNLNPNPTTTLTLTHPHHPOT 2940
Db      2881  QOARNTVDYATTTQTLHQHGVTVYIELGRDNTLTTLTHNLNPNPTTTLTLTHPHHPOT 2940
Qy      2941  HLTLNLAKTTTTHPHNTHNDNOPHTHTHLDLFTYPRONHNHYLBESTORGAGVSAAGL 3000
Db      2941  HLTLNLAKTTTTHPHNTHNDNOPHTHTHLDLFTYPRONHNHYLBESTORGAGVSAAGL 3000
Qy      3001  DPTPHPLGATLELATGCGALLAGRLSRSHPLADAVGGTVLLSGATLELATLHAGTY 3060
Db      3001  DPTPHPLGATLELATGCGALLAGRLSRSHPLADAVGGTVLLSGATLELATLHAGTY 3060
Qy      3061  VGCGRVDELTLHARLVVPDGVGVVQVAAADSEGRRLVSVVARGSGACGGGASGGVW 3120
Db      3061  VGCGRVDELTLHARLVVPDGVGVVQVAAADSEGRRLVSVVARGSGACGGGASGGVW 3120
Qy      3121  TCHASGVLEPAAAGVVNDGLAGVPRRGAVAVNDGVDRBLAGACVLPVPSGLRAVW 3180
Db      3121  TCHASGVLEPAAAGVVNDGLAGVPRRGAVAVNDGVDRBLAGACVLPVPSGLRAVW 3180
Qy      3181  RDGSDLLAEVCLPEEAWGDAAGFGALHPALLDGVVQPLSVLLPGGTGEGAGFEGVAVP 3240
Db      3181  RDGSDLLAEVCLPEEAWGDAAGFGALHPALLDGVVQPLSVLLPGGTGEGAGFEGVAVP 3240
Qy      3241  AVMGVSILHRAVGTGVAVRVSAGVGGGREAVSVVVDGAGVPVAVDRLELRFVDMGOL 3300

```

```

Db      3241  AVMGVSILHRAVGTGVAVRVSAGVGGGREAVSVVVDGAGVPVAVDRLELRFVDMGOL 3300
Qy      3301  RAVSVAGRRGSLYAVQMAEVEGVVPVCGQAMAHMEDVSGSGGPVGVVLRCPDAGAG 3360
Db      3301  RAVSVAGRRGSLYAVQMAEVEGVVPVCGQAMAHMEDVSGSGGPVGVVLRCPDAGAG 3360
Qy      3361  GGGGGGGGGGGVEVGVVGLVQVQGLERFAGSRLVVVTRGAVVAGPEDGPDVDVAGSVW 3420
Db      3361  GGGGGGGGGGGVEVGVVGLVQVQGLERFAGSRLVVVTRGAVVAGPEDGPDVDVAGSVW 3420
Qy      3421  GLVRSQAQEHPRFVLLDDTDGTCTLDLTGAGAGMVDGGRVAAVVAACEPOLAVNGERL 3480
Db      3421  GLVRSQAQEHPRFVLLDDTDGTCTLDLTGAGAGMVDGGRVAAVVAACEPOLAVNGERL 3480
Qy      3481  LAARLRLSESSGVVPVQSGDTPRARSDVPVPAQRSGVPPARSDVSGREVLPMLSGSVL 3540
Db      3481  LAARLRLSESSGVVPVQSGDTPRARSDVPVPAQRSGVPPARSDVSGREVLPMLSGSVL 3540
Qy      3541  VTGGTGVLAGAVALHLAGVGVRLDLLVSRGCPDABGABGLRAELALGAEVRIVACDVG 3600
Db      3541  VTGGTGVLAGAVALHLAGVGVRLDLLVSRGCPDABGABGLRAELALGAEVRIVACDVG 3600
Qy      3601  ERREVVRLLEGVACGPTLVGVHAAVLDATASLTPERLGTVPFAKTDALLDLBELTR 3660
Db      3601  ERREVVRLLEGVACGPTLVGVHAAVLDATASLTPERLGTVPFAKTDALLDLBELTR 3660
Qy      3661  GMELSAFVLFSSAAGILGSGAGGQNVAAANAALDALYRRRAAGLPGVSLAMGLEBSAGM 3720
Db      3661  GMELSAFVLFSSAAGILGSGAGGQNVAAANAALDALYRRRAAGLPGVSLAMGLEBSAGM 3720
Qy      3721  TGHLAGTDHRRRIIRSGLHPMSTPDALALFDALALADRVLPRADLRPAPRLDOLLP 3780
Db      3721  TGHLAGTDHRRRIIRSGLHPMSTPDALALFDALALADRVLPRADLRPAPRLDOLLP 3780
Qy      3781  ATRRRRTTRTTTGGANGAQLHARLAGOTHEOHTTLALVRSNIAVUGHTTPTDTPPD 3840
Db      3781  ATRRRRTTRTTTGGANGAQLHARLAGOTHEOHTTLALVRSNIAVUGHTTPTDTPPD 3840
Qy      3841  RAPRDGFSLTAVELRNRLSRTTGLRLPTTLAFDHPNPPTTLTHNLHTLOQOPDNAVAP 3900
Db      3841  RAPRDGFSLTAVELRNRLSRTTGLRLPTTLAFDHPNPPTTLTHNLHTLOQOPDNAVAP 3900
Qy      3901  VLAELDKESALSALDKTDSASERVTLRLKSLMRNNAPOHPRAESADDEKTSATEAZ 3960
Db      3901  VLAELDKESALSALDKTDSASERVTLRLKSLMRNNAPOHPRAESADDEKTSATEAZ 3960
Qy      3961  IFFKINDNLGLS 3972
Db      3961  IFFKINDNLGLS 3972

RESULT 2
ID 09F7T9 PRELIMINARY; PRT; 3626 AA.
AC 09F7T9;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Avermectin polyketide synthase (fragment).
OS Streptomyces avermectilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetales; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC31271;
RA Hong Y.-S., Lee J.J.;
RT "Targeted Gene Disruption of the avermectin O-methyltransferase gene
RL Submitted (JUN-2000) to the EMBL/GenBank/DBS databases.
DR EMBL; AF275943; AAG09812.1;
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.

```

DR InterPro; IPR003880; Pantne, attach.
 DR Pfam; PF00698; Acyl transf.; 3.
 DR Pfam; PF00109; ketocyl-synt.; 2.
 DR Pfam; PF02801; ketocyl-synt_C; 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR PROSITE; PS00075; ACY DOMAIN; 2.
 DR PROSITE; PS00606; B. KETOACYL SYNTHASE; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 FT NON TER 3626 3626
 SO SEQUENCE 3626 AA; 380557 MW; 6272F5F086C1A8D0 CRC64;

Query Match 84.7%; Score 17634.5; DB 2; Length 3626;
 Best Local Similarity 92.6%; Pred. No. 0;
 Matches 3419; Conservative 33; Mismatches 115; Indels 125; Gaps 21;

QY 20 VADADGGVVFVFGGQGPQWPMKRELLDASDVRESVRACEAPAPVPMDSVQVLRDS 79
 DB 1 MADBADGGVVFVFGGQGPQWPMKRELLDASDVRESVRACEAPAPVPMDSVQVLRDS 60
 QY 80 PDAGGLRDVDVQPTLFAVMISLAALWRSQGVPCAVLGHSLGEIAAAHVSGLADAA 139
 DB 61 PDAGGLRDVDVQPTLFAVMISLAALWRSQGVPCAVLGHSLGEIAAAHVSGLADAA 120
 QY 140 RVVTLMSQQTLLAGTGAALVSVAATPDELLPRIAPWTEEDNPARLAVAAVNGRSTVSGA 199
 DB 121 RVGDAWSQQTLLAGTGAALVSVAATPDELLPRIAPWTEEDNPARLAVAAVNGRSTVSGA 180
 QY 200 REAADVADLTAAOVTRMTLVDPVPAHSPLMTVAIEERVSGLLPITPRPSRIEPPHSVT 259
 DB 181 REAADVADLTAAOVTRMTLVDPVPAHSPLMTVAIEERVSGLLPITPRPSRIEPPHSVT 240
 QY 260 GGRLDTEELDAAYWRNMSSTVREPARALLQOGPKTFVMSPPHVLTMLOELADPLG 319
 DB 241 GGRLDTEELDAAYWRNMSSTVREPARALLQOGPKTFVMSPPHVLTMLOELADPLG 300
 QY 320 DTGTADTVIWTGLRRGGTLDHPLTSLAQIRGGETSATTVLSARLTALSPTQOQSLIL 379
 DB 301 DTGTADTVIWTGLRRGGTLDHPLTSLAQIRGGETSATTVLSARLTALSPTQOQSLIL 360
 QY 380 DLVAHMTAVLNDGNERTASDAPBSASFHLGFDVSWGVLELRRLSKATLRLPYTLIF 439
 DB 361 DLVAHMTAVLNDGNERTASDAPBSASFHLGFDVSWGVLELRRLSKATLRLPYTLIL 420
 QY 440 DHTPAVAARLRTAALGHLEDETPAPVPSPSGSGTAAADPPIAIGMACRFPGVSRP 499
 DB 421 DHTPAVAARLRTAALGHLEDETPAPVPSPSGSGTAAADPPIAIGMACRFPGVSRP 480
 QY 500 KDILWELAASGDALGPPTDRGWPTEQRHADPTOPGTFFYPQGGGFLHDAHPDAGFGI 559
 DB 481 KDILWELAASGDALGPPTDRGWPTEQRHADPTOPGTFFYPQGGGFLHDAHPDAGFGI 540
 QY 560 SPRRLAMDPOQRLLLETSMWAFERAGIDPLSVGSRGVAGALSPYGRMTTASSEG 619
 DB 541 SPRRLAMDPOQRLLLETSMWAFERAGIDPLSVGSRGVAGALSPYGRMTTASSEG 600
 QY 620 AADVEGHILITGTSVLSGRILAYSGLEGPAITVDTGCSASLVTLHACOSLRSGECTLA 679
 DB 601 AADVEGHILITGTSVLSGRILAYSGLEGPAITVDTGCSASLVTLHACOSLRSGECTFA 660
 QY 680 LAGGVSWSTLGMFIEFSRORGLSVDRCKAYSAADGTGMEGVMLVERLSDAVRLG 739
 DB 661 LAGGVSDVHR-GMFIERSRQCGLSVDRCCKAVSAADGTGMEGVMLVERLSDAVRLG 706
 QY 740 HRVLAUVRGSAVNDGASNGILTAIPNGPAQERVIQOALANAGLSAADVAVVEGHGTGTLG 799
 DB 707 ---LAUVRGSAVNDGASNGILTAIPNGPAQERVIQOALANAGLSAADVAVVEGHGTGTLG 763
 QY 800 DPTEAQLLATYGRAGDRPLMLGSLKSNIGHTMAAGVGIVIMVVALRGVLPRTLHV 859
 DB 764 DPTEAQLLATYGRAGDRPLMLGSLKSNIGHTMAAGVGIVIMVVALRGVLPRTLHV 823
 QY 860 DKPSPOVDWSAGAVRLLEAVPMPGDAAGRLRRAGVSGFGIGTNHVIIEAPAGGCV 919
 DB 840 DKPSPOVDWSAGAVRLLEAVPMPGDAAGRLRRAGVSGFGIGTNHVIIEAPAGGCV 819

DB 824 DEBSPQ-GLAAGAVRLLEAVPMPGDAAGRLRRAGVSGFGIGTNHVIIEAPAGGCV 882
 QY 920 AGGVLEAGAPGLAISVAESVAAPVAAPVAESVPVVPVVPVPSARSEAGLRQAQA 979
 DB 883 AGGVLEAGAPGLAISVAESVAAPVAAPVAESVPVVPVVPVPSARSEAGLRQAQA 942
 QY 980 LROYAVRPVDSLADVAGAGLACRAVLEHRAVLA-- 1015
 DB 943 LROYAVRPVDSLADVAGAGLACRAVLEHRAVLA-- 1002
 QY 1016 DRELVGGLAALAGEDRVTTGHPAGRGGVVFVFGGQGPQWPMKRELLDASDVRESVRACEAPAPVPMDSVQVLRDS 1075
 DB 1003 PARTSARIGALAAAGSSAALTGHPAGRGGVVFVFGGQGPQWPMKRELLDASDVRESVRACEAPAPVPMDSVQVLRDS 1062
 QY 1076 RMOACEALAPWDMVVDVILRRDGDAAWEBAVDVQPLFSSVMISLAALMSYGIERN 1135
 DB 1063 RMOACEALAPWDMVVDVILRRDGDAAWEBAVDVQPLFSSVMISLAALMSYGIERN 1122
 QY 1136 AVLGSGEITAAHVCGALSLKDAKTVALRSBALAVRSGGMAVPLPAQVEQLIGE 1195
 DB 1123 EVLGSHKDEITAAHVCGALSLKDAKTVA-----LPPQVEQLIGE 1163
 QY 1196 RMAGRLWVAAVNGRSTAVSGDAVDEVLAAYCAGTVARARIPVDYASCHPVQPLREE 1255
 DB 1164 R-GGRILWVAAVNGRSTAVSGDAVDEVLAAYCAGTVAR-RIPVDYASCHPVQPLREE 1221
 QY 1256 LLELGDISPQSGVPFESTVEGTWLDTTLDAAWYRNLRHOPVRSFSDAVALADQHRV 1315
 DB 1222 LLELGDISPQSGVPFESTVEGTWLDTTLDAAWYRNLRHOPVRSFSDAVALADQHRV 1281
 QY 1316 FVEVSPHPLVPAIEDTTEDETAEDVTAIGSLRGDNDTRRFLTALAHTHTGTGTPPTWH 1375
 DB 1282 LLEVSPHPLVPAIEDTTEDETA-----R-----HCDRQPRPRER 1318
 QY 1376 H-----HYTH---HTHP---HPTHLDLPT-----YPRQHQYTHLESSQPGA 1412
 DB 1319 HPLLPORLRDLPYRRHQTHNVAPLDPRHLPPLDAPRARI.PQOHYHLESSQPGA 1378
 QY 1413 GSGSGAGAGSGAGSGRAGTAGTAESRWDVAASODLETVAITLAVPSAGLDTVPA 1472
 DB 1379 GSGSGAGAGSGAGSGRAGTAGTAESRWDVAASODLETVAITLAVPSAGLDTVPA 1438
 QY 1473 LSAWHRHQDQARINWTQYETKWP.LLPTTHOPHOTWILAIPEOTQHPIINILTNLH 1532
 DB 1439 LSAWHRHQDQARINWTQYETKWP.LLPTTHOPHOTWILAIPEOTQHPIINILTNLH 1498
 QY 1533 HHGITPRLPLTNHTNPNQHLNHTLNTRQOANTTGATGLLSLADDETHPHNPT 1592
 DB 1499 HHGITPRLPLTNHTNPNQHLNHTLNTRQOANTTGATGLLSLADDETHPHNPT 1558
 QY 1593 PTGTLNLTLQHTNTPRPLMYATNAATTNHPNDPLNHPQOATWGLARTLLEHPT 1652
 DB 1559 PTGTLNLTLQHTNTPRPLMYATNAATTNHPNDPLNHPQOATWGLARTLLEHPT 1618
 QY 1653 HTAGIDLPPTPHLLOHLQTLTOHPNQTOLAITGTGHTNRRLPTTLPTPHNPTPT 1712
 DB 1619 HTAGIDLPPTPHLLOHLQTLTOHPNQTOLAITGTGHTNRRLPTTLPTPHNPTPT 1678
 QY 1713 PHGTTILTGAGALNHLNHTLTHPTOHLNLSRTGPHTPHONHTLOOQGHILTI 1772
 DB 1679 PHGTTILTGAGALNHLNHTLTHPTOHLNLSRTGPHTPHONHTLOOQGHILTI 1738
 QY 1773 TTCDDTSNPDOL-QOLNLTIPROGRPLTVIHTAGILDDATLNTLPTQOLNNVLAQKASHAH 1831
 DB 1739 TTCDDTSNPDOL-QOLNLTIPROGRPLTVIHTAGILDDATLNTLPTQOLNNVLAQKASHAH 1798
 QY 1832 LHLQLOHTPLTAFLVYSSAAATFGAPGOANYAANAAYDALAHNHTNHLPATSLAMGT 1891
 DB 1799 LHLQLOHTPLTAFLVYSSAAATFGAPGOANYAANAAYDALAHNHTNHLPATSLAMGT 1858
 QY 1892 WQNGNLADSDKADAYLDRGRFMSBELATAVTOAIADTERYVVIADIDMSKIEHTSQ 1951
 DB 1859 WQNGNLADSDKADAYLDRGRFMSBELATAVTOAIADTERYVVIADIDMSKIEHTSQ 1918

QY 1952 TSDELVSAREREPAPQPTPEAEHLKTLAQTSAADQPAALLELVRDHVAALRHADKAI 2011
 DB 1919 TSDLSAAREREPAPQPTPEAEHLKTLAQTSAADQPAALLELVRDHVAALRHADKAI 1978
 QY 2012 APDQSPALGPDSITAVEFNNLLIKATGLRPLVSLVDHPPTAKLAVHLONOLRGTAES 2071
 DB 1979 APDQSPALGPDSITAVEFNNLLIKATGLRPLVSLVDHPPTAKLAVHLONOLRGTAES 2038
 QY 2072 APSAAATASAVTEPITAIVGACRPFQVTSADPFMDLISSEBDAIGSFPTORGMPLDT 2131
 DB 2039 APSAAATASAVTEPITAIVGACRPFQVTSADPFMDLISSEBDAIGSFPTORGMPLDT 2098
 QY 2132 LYDPDPHPCGTCTYRNCGFLYDAGHPDAEFPGISPREALAMDPOORLLLETAMETIEHAG 2191
 DB 2099 LYDPDPHPCGTCTYRNCGFLYDAGHPDAEFPGISPREALAMDPOORLLLETAMETIEHAG 2158
 QY 2192 INPHTLHGTPTGVTNGODYALRVANAGOSTDGPALTGTAGSVISGRISYTFGFECPA 2251
 DB 2159 INPHTLHGTPTGVTNGODYALRVANAGOSTDGPALTGTAGSVISGRISYTFGFECPA 2218
 QY 2252 VSVDTACSSLVMLHLACALRAGECSMALAGVTWSSPGAFVFSRQGLAADGCKA 2311
 DB 2219 VSVDTACSSLVMLHLACALRAGECSMALAGVTWSSPGAFVFSRQGLAADGCKA 2278
 QY 2312 FSAAADGTGMEGVGMLLVERLSDAHNRHVRVLAVGSANODGANGLTAPNGPQOR 2371
 DB 2279 FSAAADGTGMEGVGMLLVERLSDAHNRHVRVLAVGSANODGANGLTAPNGPQOR 2338
 QY 2372 VIRQALNAGLSAGDVDAVEAHGTTIGDPIEAQALLATYGGDRAGEGRLMGSVKSXN 2431
 DB 2339 VIRQALNAGLSAGDVDAVEAHGTTIGDPIEAQALLATYGGDRAGEGRLMGSVKSXN 2398
 QY 2432 GHQQAAGVAGVUIMMVALRHGILPRILYHDEBPHNDMSAGANQOLLTEVPMPGGGRL 2491
 DB 2399 GHQQAAGVAGVUIMMVALRHGILPRILYHDEBPHNDMSAGANQOLLTEVPMPGGGRL 2458
 QY 2492 RRAQVSGFVSGTNAHYILEBAPADVPGRPAGEGSDGDEAAAGSPGVMPVLVAKS 2551
 DB 2459 RRAQVSGFVSGTNAHYILEBAPADVPGRPAGEGSDGDEAAAGSPGVMPVLVAKS 2518
 QY 2552 QPALRAQALNAHLTDHPGLDLADVGYTLAHABAVFDHBRATLLAADRTPLQALQALAA 2611
 DB 2519 QPALRAQALNAHLTDHPGLDLADVGYTLAHABAVFDHBRATLLAADRTPLQALQALAA 2578
 QY 2612 GERPRATIHSSAPGSGTGERAGKTAIFCSGGQTORGMAHGLYHTRPVFAALNDICTH 2671
 DB 2579 GERPRATIHSSAPGSGTGERAGKTAIFCSGGQTORGMAHGLYHTRPVFAALNDICTH 2638
 QY 2672 LDRPLDRPLRLTQONDNEDBAALAAQOTRYAOPALFAFOVALHRLLTGCGYHTRPYUA 2731
 DB 2639 LDRPLDRPLRLTQONDNEDBAALAAQOTRYAOPALFAFOVALHRLLTGCGYHTRPYUA 2698
 QY 2732 GHSIGBITAANLAGILTLDTATTLITQORATLMQMPRGMTTLHTTHHITHLTAHEND 2791
 DB 2699 GHSIGBITAANLAGILTLDTATTLITQORATLMQMPRGMTTLHTTHHITHLTAHEND 2758
 QY 2792 LATAAITPTSLVSGRPHYQATTTTCCOOGIKTKLPTNNAHSHNTPRILNQLHQT 2851
 DB 2759 LATAAITPTSLVSGRPHYQATTTTCCOOGIKTKLPTNNAHSHNTPRILNQLHQT 2818
 QY 2852 QTLTYRPHRPLTANTPRDOLLRNHYMTOOARNTVUAATTTQCHONGVTTYELAPDN 2911
 DB 2819 QTLTYRPHRPLTANTPRDOLLRNHYMTOOARNTVUAATTTQCHONGVTTYELAPDN 2878
 QY 2912 TLTLTLTHNHLNRPRTTLTLTLTHNHNQTHLLTNLAKTITTHNHNHNDONPHYTH- 2970
 DB 2879 TLTLTLTHNHLNRPRTTLTLTLTHNHNQTHLLTNLAKTITTHNHNHNDONPHYTH- 2937
 QY 2971 LDRPTYRPHNHNHVLBSTQBGAGVSAAGDPTPHRLGATLELATGSGALLAGRLSLRS 3030
 DB 2938 LDRPTYRPHNHNHVLBSTQBGAGVSAAGDPTPHRLGATLELATGSGALLAGRLSLRS 2997

QY 3031 HPMLADHAGVGTLLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPVGVSVQVGA 3090
 DB 2998 HPMLADHAGVGTLLLSGATFLELALHAGTYVGCDDRVDELTLHAPLVVPVGVSVQVGA 3057
 QY 3091 AADGEGRLVSVYARSGSACGGGAGAGCTWTCHASGVLVEAAGGVVVDGLAGWPPRGA 3150
 DB 3058 AADGEGRLVSVYARSGSACGGGAGAGCTWTCHASGVLVEAAGGVVVDGLAGWPPRGA 3117
 QY 3151 VAVDVGVDRDLAAGCCTVGPVSGLRVAVRDGDDLAEVCLPEBAMGDAAFGHLHALL 3210
 DB 3118 VAVDVGVDRDLAAGCCTVGPVSGLRVAVRDGDDLAEVCLPEBAMGDAAFGHLHALL 3177
 QY 3211 DGVVQPLSVLLPGCTGFGEBGAGFGBGVRYPAVWGVSLRHAGYTGVRVSAVGRGGRE 3270
 DB 3178 NGVVGQPLSVLLPGCTGFGEBGAGFGBGVRYPAVWGVSLRHAGYTGVRVSAVGRGGRE 3237
 QY 3271 AVSVVVGDEAGVAVSVDRLELRPVDMGOLRAVSVSAGRRGSLYAVQMAEVGPVCCGA 3330
 DB 3238 AVSVVVGDEAGVAVSVDRLELRPVDMGOLRAVSVSAGRRGSLYAVQMAEVGPVCCGA 3297
 QY 3331 WAMHEDVSGSGGPVPVGVVTLRCPDAGAGGGGGGGGGVGEVVGVLGVVQCMGLGERF 3390
 DB 3298 WAMHEDVSGSGGPVPVGVVTLRCPDAGAGGGGGGGGGVGEVVGVLGVVQCMGLGERF 3357
 QY 3391 AGSRLVVVTRGAVVAGPEDEGPVUVVGSVGLVRSQAEBHDFVLLDLDTDGTDLDTG 3450
 DB 3358 AGSRLVVVTRGAVVAGPEDEGPVUVVGSVGLVRSQAEBHDFVLLDLDTDGTDLDTG 3417
 QY 3451 AGAGMGVDGGRVAAVAVACBEPOLAVRGERLLAARLKRLEBSGDVPAPQSGDTPARRSDVP 3510
 DB 3418 AGAGMGVDGGRVAAVAVACBEPOLAVRGERLLAARLKRLEBSGDVPAPQSGDTPARRSDVP 3477
 QY 3511 AQRSGGVPARRSVDSVGREVLPMLSGGSVLYVGGTGVLAGAIVRHLAGVGVVDLLVSR 3570
 DB 3478 AQRSGGVPARRSVDSVGREVLPMLSGGSVLYVGGTGVLAGAIVRHLAGVGVVDLLVSR 3537
 QY 3571 RGPDAFGABGLRAELALGAENRIVACDVGERRERVRLEGVAGCPPLTGVAAGAVALD 3630
 DB 3538 RGPDAFGABGLRAELALGAENRIVACDVGERRERVRLEGVAGCPPLTGVAAGAVALD 3588
 QY 3631 ATTASLTPERLGTVFAKVDAAL--LIDELT 3659
 DB 3589 SAVASVT-VRLPVASDVKREKAAVAATVEEFT 3619

RESULT 3
 Q9SOR4
 ID Q9SOR4 PRELIMINARY; PRT; 5532 AA.
 AC Q9SOR4;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 DE Type I polyketide synthase AVES 3.
 GN AVEA3.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 NC NCBI_taxid=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9380548; PubMed=10449723;
 RA Ikeda H., Nonomiyama T., Usami M., Ohta T., Omura S.;
 RT "organization of the biosynthetic gene cluster for the polyketide
 RT avermectin macrocyclic avermectin in Streptomyces avermitilis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514 (1999).
 DR EMBL; AB032367; BAA84478.1; -;
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; prantne attach.
 DR InterPro; IPR000834; Zn_cardopect.
 DR Pfam; PF00698; Acyl_transf. 3.
 DR Pfam; PF01009; ketoacyl-synt. 3.
 DR Pfam; PF02801; ketoacyl-synt. C; 3.

Db 1976 RSGQTGVFAGLIPQAVGRPHENAAADTEGVLTIGTSGVASGRISTYTFEGEGRPAVSDVT 2035
 Qy 2257 ACSSSLVALHLACQALRAGECSMALAGVTVMSSPGAFVEFSPRORGLAADGCKAFSA 2316
 Db 2036 ACSSSLVALHLACQALRAGECSMALAGVTVMSSPGAFVEFSPRORGLAADGCKAFSA 2095
 Qy 2317 DGTGMEGVGMILLVERLSDAHRNGHRLAVVRGSAVNQDASNGCLTAPNGPQQRVITRO 2376
 Db 2096 DGTGMEGVGMILLVERLSDAHRNGHRLAVVRGSAVNQDASNGCLTAPNGPQQRVITRO 2155
 Qy 2377 LANAGSAGVDVAVEAHGTGTGDPLEAOLALLATYQODRAGCGPLMGSVKSNVHTQA 2436
 Db 2156 LANAGSAGVDVAVEAHGTGTGDPLEAOLALLATYQODRAGCGPLMGSVKSNVHTQA 2215
 Qy 2437 AAGVAGYIKVMALRHGLPRTLHVDESPHVMDSAGAVOLLTETVWMPGEGRLRAGV 2496
 Db 2216 AAGVAGYIKVMALRHGLPRTLHVDESPHVMDSAGAVOLLTETVWMPGEGRLRAGV 2275
 Qy 2497 SSGVSGTNAHVILEA-----PADDPGPGPAGDAGSDDEAAAGSPGVPMVLV 2547
 Db 2276 SSGVSGTNAHVILEAPAHNIPSDTPADDPGPGPAGDAGSDDEAAAGSPGVPMVLV 2334
 Qy 2548 SASOPRLAQAOLAHNLTDHPGLADVGYTLAARAVPDRATITLADPDTFLQALQ 2607
 Db 2335 SASOPRLAQAOLAHNLTDHPGLADVGYTLAARAVPDRATITLADPDTFLQALQ 2394
 Qy 2608 ALAAGEPRPAVHSSAFGTGTGEAAGKTAFTCSGGGTORPGMAHGLYHTHPVPAALND 2667
 Db 2395 ALAAGEPRPAVHSSAFGTGTGEAAGKTAFTCSGGGTORPGMAHGLYHTHPVPAALND 2454
 Qy 2668 ICTHLDPHLDHRLPRLYLTQNDNDN-----EDAAALQOTRYAOPALFAFOVALHRLTLDG 2722
 Db 2455 ICTHLDPHLDHRLPRLYLTQNDNDN-----EDAAALQOTRYAOPALFAFOVALHRLTLDG 2514
 Qy 2723 YHTTPHYAGHSIGETAAHLAGILITTDATYTLTGATLMQMPRGTMTTLTTPHNT 2782
 Db 2515 YHTTPHYAGHSIGETAAHLAGILITTDATYTLTGATLMQMPRGTMTTLTTPHNT 2574
 Qy 2783 HHHTAHENDALAIANTPTSLVIGSPHTYOHITTLTCCQOOGIKTKLPTNHAHSPHTNP 2842
 Db 2575 HHHTAHENDALAIANTPTSLVIGSPHTYOHITTLTCCQOOGIKTKLPTNHAHSPHTNP 2634
 Qy 2843 ILNQLQHOTQTLTYNRPHTLITANTPRDQLTTPHWTQOARVTVDAITTLQHLQNGVT 2902
 Db 2635 ILNQLQHOTQTLTYNRPHTLITANTPRDQLTTPHWTQOARVTVDAITTLQHLQNGVT 2694
 Qy 2903 TYIELGPDNTLTTLTHNLNRPRTTTLTLPHNHPTQHLTLIAKTTTWHPHNTYHND 2962
 Db 2695 TYIELGPDNTLTTLTHNLNRPRTTTLTLPHNHPTQHLTLIAKTTTWHPHNTYHND 2754
 Qy 2963 NQHTHTHLDLPYRPHNHVLESTQPGANVSAAGLDPTERPLGATLELATDQCAL 3022
 Db 2755 NQHTHTHLDLPYRPHNHVLESTQPGANVSAAGLDPTERPLGATLELATDQCAL 2814
 Qy 3023 AGRLSLRSHWMLDHAAGTTLVLSGATFLELALNAGTYVGCDDRELTLLAPLVPRVDCG 3082
 Db 2815 TGRLSLRTHWMLAGHAGVLLRGTAFALALHAGSVCDDRELTLLAPLVPRVDCG 2874
 Qy 3083 VSVQVGAALADGGRRLVSVYARAGSAGCGGASGCVTCHASGVLEAAAGVV-VDLG 3141
 Db 2875 VTLQVVAIAARDGSRMTIHSRGECSGADASASAMTAAAGVLSRAPDDTASVEL 2934
 Qy 3142 AGWPRPAGAVAVDVGRDLRAGACVLGRVFGSLRABVMDGGDLAEVCLPREAWMDAA 3201
 Db 2935 AGWPRPAGATPRVDLNTAYDQMDADGRPYGALFQGLRAMMYGDDILEARLPEVSGDAA 2994
 Qy 3202 GFGHLRALDGVVQPLSVLLPGSTFGFEGAGFEGGVAVPVWVGSLHRAVGTGVRYVS 3261
 Db 2995 AYGLHRLDLDAALQ-----GTGLLSVAGPGRPV-VPHVWGLRRTTHGAVSVARACLS 3045
 Qy 3262 AVRGGRRAVSVVGVDEAGVPVAVVDRLRLPYDMQOLRAVSVSAGRGSLLYAVQMAEV 3321
 Db 3046 TLGATGA--AVCVRLTDTGTGVPAVSVDRLELRPMQOLRAVSVSAGRGSLLYAVQMAEV 3103

Qy 3322 GPVPVCGOAMAMHEDVSGSGGFPVPGVVLRCPDAGAGGGGGGGGVGVGVVLGVV 3381
 Db 3104 GPVPVCGOAMAMHEDVSGSGGFPVPGVVLRCPDAGAGGGGGGGGVGVGVVLGVV 3159
 Qy 3382 QGWLGLERFAGSRLLVVTVTGAVVAGPEDPVDVVGASVNGVYLSAOAEHPDRVLLDLO 3441
 Db 3160 QGWLGLERFAGSRLLVVTVTGAVVAGPEDPVDVVGASVNGVYLSAOAEHPDRVLLDLO 3219
 Qy 3442 --DTGTDLDTG--AGAGKGVDDGVRVAVVAACBPOLAVGERLLARLKRSSGCVPAQ 3497
 Db 3220 DLDSGADADAGNAGMSGSLDGGVRVAVVAACBPOLAVGERLLARLKRSSGCVPAQ 3272
 Qy 3498 RSGDTRARRSDVPAQSGGVPARSVDSVGRREVLPMLSGSVLYVTGGTGVLAGAVALRH 3557
 Db 3273 -----PDVSGREVLPHLSGSGSVLYVTGGTGVLAGAVALRH 3308
 Qy 3558 GVCVGRDLLVSRGDPAGAEGLRAELALGAERVIVACDVGERRVRLLEGVPAGCP 3617
 Db 3309 GVCVGRDLLVSRGDPAGAEGLRAELALGAERVIVACDVGERRVRLLEGVPAGCP 3368
 Qy 3618 LTVGVHAAGVLDATIASLTPERLGTVPAAKVDALLDLBELTGMELSAFVLFSSAAGIL 3677
 Db 3369 LTVGVHAAGVLDATIASLTPERLGTVPAAKVDALLDLBELTGMELSAFVLFSSAAGIL 3428
 Qy 3678 GSAGQGNVAAANALDALAVRRRAAGLPGVSLAMGLMESAGMTGHLAOTDHRRIIRSL 3737
 Db 3429 GSAGQGNVAAANALDALAVRRRAAGLPGVSLAMGLMESAGMTGHLAOTDHRRIIRSL 3488
 Qy 3738 HPWSTPDALAFDALALADRLPVLPRADLRPAPELPLDOLPATRRRTTITTTGGADN 3797
 Db 3489 HPWSTPDALAFDALALADRLPVLPRADLRPAPELPLDOLPATRRRTTITTTGGADN 3548
 Qy 3798 GAOLHARLAGQTHEOQHTLLLVRSHTATVUGHTTPTDITPPRARRDLGFDLTVAE 3857
 Db 3549 GAOLHARLAGQTHEOQHTLLLVRSHTATVUGHTTPTDITPPRARRDLGFDLTVAE 3608
 Qy 3858 NRLSRTTGLRLPTTLAFDHPNPPTTLTHNLHTOL 3890
 Db 3609 NRGENTGLRLPTSLVFPDQPMATTLARHLREL 3641
 RESULT 4
 Q9SOR7 PRELIMINARY; PRT; 6239 AA.
 AC Q9SOR7;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
 DE Type I polyketide synthase AVES 2.
 GN AVER2.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N. A.
 RX MEDLINE=93380548; PubMed=10449723;
 RA Ikeda H., Nonomiya T., Usami M., Ohta T., Omura S.;
 RT "Organization of the biosynthetic gene cluster for the polyketide
 RT anthelmintic macroide avermectin in Streptomyces avermitilis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9509-9514 (1999).
 DR EMBL AB032367; BAA84475.1; -
 DR InterPro: IPR001227; Ac_Transferase.
 DR InterPro: IPR000794; ketoacyl-lynt.
 DR InterPro: IPR003880; peptidyl-attach.
 DR InterPro: IPR000169; SHprot acetate.
 DR InterPro: IPR000834; Zn carbopapt.
 DR Pfam: PF006698; Acyl_transf. 4.
 DR Pfam: PF00109; ketoacyl-lynt; 4.
 DR Pfam: PF02801; ketoacyl-lynt; C; 4.
 DR Pfam: PF00550; pp-binding; 4.
 DR PROSITE, PSS0075; ACP_DOMAIN; 4.

DR PROSITE; PS00606; B KETOACYL SYNTHASE; 4.
 DR PROSITE; PS00133; CARBOXYPEPT ZN 2; UNKNOWN_4.
 DR PROSITE; PS00012; PHOSPHOPANTHETINE; 3.
 DR PROSITE; PS00639; THIOL PROTEASE_HIS; UNKNOWN_1.
 KW Phosphopantetheine; Transferrase.
 SQ SEQUENCE 6239 AA; 666283 MW; A706280940B502FA CRC64;

Query Match 54.4%; Score 11330.5; DB 2; Length 6239;
 Best Local Similarity 44.3%; Pred. No. 0;
 Matches 2512; Conservative 312; Mismatches 961; Indels 1683; Gaps 61;

QY 1 VQRMDSGEPRPA-----AGEVLGVADEADGVVFVPGQGFQWFGMGRELLDASDF 53
 DB 565 LQALAAAG-BEPHVAIHSSAPGCTGTG--EAAAGTAFICSGQGTQGRGMAGLHTHPVF 620
 QY 54 RESVRACEAARFAYVDMDSVEQVLRDSPD-----ARGLDVNDVQPTLFRVMTSL- 102
 DB 621 AALANDICTHLDHPLRLPLLTQDPNTQDPTTLEBAALLQOTRYAQRLEFAFQVALH 680
 QY 103 AALMRSGVEPCAVLGHSLSGEIAAAHVSGSLDAARVVTLMGSOAQTTLAAGTALVSA 162
 DB 681 RLITDGHITPRHYVAGHSLSGEITAAHLGITLTDATLITQAKTLMQTPR-PGTMTTLH 739
 QY 163 ATPDELLPRIAPTEDNPARLAVAANVGPRSTVVGAREAVADLVADLTAAQVETRMIPV 222
 DB 740 TTPHITHTLTAHEND---LAIATAINTPSLVISGTPHVOHITTLCCQOQGIKITLPLT 795
 QY 223 DVRAHSLMTAIEERVVSGLLPIPRPSKIPFHSVVGRLDPTRELDAAWYRMSSTVR 282
 DB 796 NHAHSHHTNPRLNQLHQHTQTLTYNRPHTPL---ITANTRPDLLTPHWTQOARNTVD 852
 QY 283 FEPAARLLLOQGRFTEVMSRHPVLTWGLQELABDLDTTGTADVTIMGLTRGQGLYLDH 342
 DB 853 YATTTQHLHQGVYTYIELGRDNLTLTLTDNLPLENTFTTIT-----LTHPHNHQTH 905
 QY 343 FLTSLAO-----LRGHE----- 355
 DB 906 LITLTLAKTTTTHNHTHNNQPHHTHLDLPTYPFQHNHYWLQPGKESDPSSEGRE 965
 QY 356 --TSATTVLSARLTALSPTQOQSLLDLVRAHTMAVLNDQNETTADAPRSASFALIFG 413
 DB 966 QATTPSPRLDVLVGSKRQERDEBELRLVHTHAAVL---GHAETREIVYNNKAFKELGF 1021
 QY 414 DSVNGVELRNRLSKATGLRLPVTLLIFDHTTPAAVAARLTALGHLDEDTAPVDSPSGH 473
 DB 1022 DSLAAILDRRLADVDLRLPATLIFDYPTPMALCGRLRAALYGADGTTTRLPLT----- 1077
 QY 474 GGTAAADDPALIGMACRPFQGVRSRYDLWELAASGDALIGRFTDRGMPTBQRHAODPT 533
 DB 1078 --AVPADEPIAIVGMACRYPGDVTVDDLMQVVGSGHDALIGRFTNRGMDLDTLYNDRD 1135
 QY 534 QPGTFYQGGGFLDAAHFDAGFGCISPREALANDPOORLLLETSMFAFERAGIDPLSVR 593
 DB 1136 HHGTSYRSRSGFPLVDAGNFDPDFPGISPREALANDPOORLLLETSMFSTHACINPDSLR 1195
 QY 594 GSRTYFAGALSFQYGRMDTASSEGAADVEGHILTTGTSVLGRSLAYFGLGEPATIV 653
 DB 1196 GTPRGVRAGLTYNDYARFPTA-----PAGFEGYLGHSAGSISAGRAVYALGLEGRPLTV 1251
 QY 654 DTGCSASLVTLHLACQSLRSGETLALAGVSVNSTLGMPLIEFSRQRLGSLVDGCKAYSA 713
 DB 1252 DTGCSASLVTLHLACQSLRSGETLALAGVSVNSTLGMPLIEFSRQRLGSLVDGCKAYSA 713
 QY 714 AADGTGGEVGMILVRLSDAVALRGRVLAUVRGSAVNODGASNGLTAPNGRQOEVR 773
 DB 1312 AADGTGGEVGMILVRLSDAVALRGRVLAUVRGSAVNODGASNGLTAPNGRQOEVR 1371
 QY 774 QALANAGLVAADVVEGHGTGTLGDPRIEAOALLATYGO--RADDRLMILGSLKSNIGHT 832
 DB 1372 LALANALUTPADVDAVEAHGCTTLGDPRIEAOALLATYGO--RADDRLMILGSLKSNIGHT 832
 QY 833 MAAAGVGAVIKMVALEGVLPRTLLAVDKSPQVDMSSAGAVRLITTEAVPMFGDAAGELRR 892

DB 1432 QAAAGVGAVIKMVALENGILPRTLLHVBESPHVDMSSAGAVOLLTETVMPG-GEGRLLRR 1490
 QY 893 AGVSSFGIGGTNAHVILEEAPAGGCVAGGVLEGAQGLAISVAESVAAPVA----- 944
 DB 1491 AGVSSFGVSGTNAHVILEEAPAIN--IPSTTPADDDAPGEA--AADVPGEAADDGAGTG 1546
 QY 945 -VASAPVAESVPEVPEVPPVVSARSSEAGLRAQAEALROYAVRPPVSLADVAGLACGR 1003
 DB 1547 EATGPAAGS-----PGWPMVLVSAGKQOPALRAQALNHLTDHPGLDLADVGTLLAHAR 1601
 QY 1004 AVLEHRAVLAAREELVQGLALMAGEPRKRVTTGHAARGD-----RGVVVFPFGGG 1058
 DB 1602 AVPDHRAVTLAARDPFLQALQALAGEBPRAVHSSAPGCTGTGAAGKTAFCISGQGT 1661
 QY 1059 QWAGMGVRLASPVFAARWQACEBALPMDVSVVDILRRAGD-----AVWR 1108
 DB 1662 QRPGMAGLHTHPVRAALANDICTHLDHPLRLPLLTQDPNTQDPTTLEBAALLQO 1721
 QY 1109 ADVVQPVLSVWVSL-AALMRSGIEPDAVLGHSOGEIAAAHVCSALSKDAKTVALRS 1167
 DB 1722 TVPAQALFAFQVALRHLITDGHITPRHYVAGHSLSGEITAAHLGITLTDATLITQRA 1781
 QY 1168 RALAAYRGCSMAVSPRPAQEVQLGERKAGLWAAVNGPRSTAVSGDAEAVDEVLAY 1227
 DB 1782 -TLMQTPRGCTMTTTLTPHHTHNLTAH-ENDLAATAINTPSLVISGTPHVOHITTL 1839
 QY 1228 CAGTVARARIPVDVASHCPVQRLREBELLELDGISPOSGVFPFSTVEGTLDTTLD 1287
 DB 1840 CQOQGIKTTLPTKMAFHSHTNPILNQLHQHTQTLTYNRPHTPL---ITANTRPDOLLT 1896
 QY 1288 AAYVRENLMQPVAFSDAVOALADGHRVFEVSRPHPLVRALEDTEDAEDVTAIGSLR 1347
 DB 1897 RHYWTQOARNTVYUATTTQHLHQGVYTYIELGRDNLTLTLTDNLPLENTFTTITLTH 1954
 QY 1348 RGNDRRRLTALAHNTTGIGCTTWN--HNTHNHTHNPHTHLDLPTYPFQHNHYWLE 1406
 DB 1955 RHNHPQTHLNTLNAK-----TTTWNHNPHTHNNQPHHTHLDLPTYPFQHNHYWLE 2007
 QY 1407 SSQPGAGSGGAGAGAGAGAGTAGTAEVESRMDAVARODLETVATTLAVPPSAGL 1466
 DB 2008 STOPGAGS-----GSGSGSGRAGTAGTNEVSRMDAVARODLETVATTLAVPPSAGL 2061
 QY 1467 DTVVPAISAMHRHQDQARINTWYQETWKRLTLPTTHOQHTWLLAIBETQHNHPHTN 1526
 DB 2062 DTVVPAISAMHRHQDQARINTWYQETWKRLTLPTTHOQHTWLLAIBETQHNHPHTN 2121
 QY 1527 ILTNLHNGITPRPLNLNHTHTNPROLHNTLHNTROAOQNTGALTGLSLALDETPH 1586
 DB 2122 ILTNLHNGITPRPLNLNHTHTNPROH---LHNTROAOQNTGALTGLSLALDETPH 2177
 QY 1587 RHNHPRTGTNLNLTLQCHTQHTNPRPLMYUATTTNAATTHPNDRPLNHPROAQWGLART 1646
 DB 2178 RHNHPRTGTNLNLTLQCHTQHTNPRPLMYUATTTNAATTHPNDRPLNHPROAQWGLART 2237
 QY 1647 LLENHPTAGIIDLPTTPRTLOHLQTLTORHNOLOLAIRTTGHTNRRLPRTLLPTN 1706
 DB 2238 LLENHPTAGIIDLPTTPRTLLNLLOLLOHNOLOLAIRTTGHTNRRLPRTLLPTN 2297
 QY 1707 ORPTTPRGHTLLITGGTALATNLTNLHTTHORTNLTSRGRVTPHNOHTTLOOK 1766
 DB 2298 ORPTTPRGHTLLITGGTALATNLTNLHTTHORTNLTSRGRVTPHNOHTTLOOK 2357
 QY 1767 GHILTTTCPTSNPDLOLOLNTIPRONPLTYVHAGLIDDTLNLPTTOLANVLRAR 1826
 DB 2358 GHILTTTCPTSNPDLOLOLNTIPRONPLTYVHAGLIDDTLNLPTTOLANVLRAR 2417
 QY 1827 AHSANHLQTLQTPPLTAFLVYSSAATGAPQOANYAANAYLDAIHNHHTHNPATS 1886
 DB 2418 AHSANHLQTLQTPPLTAFLVYSSAATGAPQOANYAANAYLDAIHNHHTHNPATS 2477
 QY 1887 IANGTQNGAGLSDAKARAYLDRGRFRMSPELATAVUQALADTERRVUVDIDWMSKI 1946
 DB 2478 IANGTQNGAGL--TGQVSEHLRRGMFAMPBELAVTAVDGALA--SGRPSLLVADIDWKKL 2535

QY	1947	EH--TSQTSDLV-----SAAEREPVQRPRTPAELHKTLANQTSADQRAALLLELVHDV	1999	QY	2987	-----	2986
Db	2536	GPVLSKSSVLEBDDLPQAQSTBEARSTVEQTESTNLRLQMLGRSRSQEBEELSLVIRIHS	2595	Db	3611	NEEDLSALETLDIDASALDVTVPALSAWHHQDOARINTWTYQETWKEPLPTTHQPH	3670
QY	2000	AAVLRLHADPRAIAPDGSFRALGSDSLTAVFERNLLKATGLRLPVSLVPRHPRAKLAHV	2059	QY	2987	-----	2986
Db	2596	AAVLGDDSDSEAIPPGRJLFRDLGFSILAVALRNHLAAQTELALPTTLVFPYPSPTXIAQF	2655	Db	3671	QTWLAIPETQTHPHRTINILTNLHHNGITPRVLTVNHTHTWBOHNLHNTLHNTROQAOMH	3730
QY	2060	LOVQLGTAESAP--SAAVTAESVTEBIAIVGMACRPPGVTASDDWDLISSRODA	2117	QY	2987	-----	2986
Db	2656	LSLEI-----AEFQPDNSTPLPRPRAELDEPIALVGMACRPGGVTASDDWDLISSRODA	2711	Db	3731	TGSPITGLLSLALDETPRHNRHPRTGTGLNLTLRQTHQTHPRPLWYATTNATTHP	3790
QY	2118	IGGFPTDRGMDLTLYDPPDHPGTCYTRNGGFLYDAGHFDAEFEGISPREALAMPQOR	2177	QY	2987	-----	2986
Db	2712	IGGFPTDRGMDLTLYDPPDHPGTCYTRNGGFLYDAGHFDAEFEGISPREALAMPQOR	2771	Db	3791	NDELTHPTQAQTFWGLARTLLLENHTHTAGIDLPTTPRTHLNLQTLTQRPHQTLAI	3850
QY	2178	LLELETAMETIEHAGINPHTLHGTPGVFTGTNGODVALRVNHAQOSTDGFALGTAGSVI	2237	QY	2991	-----	2990
Db	2772	LLELETAMETIEHAGINPHTLHGTPGVFTGTNGODVALRVNHAQOSTDGFALGTAGSVI	2831	Db	3851	RTTGTHTRBLPTTLPTTHQRPRTPRHGTTLITGSGALATHLTNLHTHQPONHLLTS	3910
QY	2238	SGRISTYFPEGEPAVSVDTCSSSLVALHLACOALRAGCSMALAGGVTWMSRPGAFVEF	2297	QY	2991	-----	2990
Db	2832	SGRISTYFPEGEPAVSVDTCSSSLVALHLACOALRAGCSMALAGGVTWMSRPGAFVEF	2891	Db	3911	RIGPHTPHAOHLTTQLQOKGHLTTTCDSNPDLQOLLNTPRQHPRLTVTHTAGVNL	3970
QY	2298	SRORGLAADGHCRAFAAADGTGMEGVGMILVERLSDAHNGHRVLAVVRGSAVNDGA	2357	QY	2991	-----	2990
Db	2892	ARORGLAADGHCRAFAAADGTGMEGVGMILVERLSDAHNGHRVLAVVRGSAVNDGA	2951	Db	3971	PAVSETDASFSSVTAAKATGAALHELLDHEHLEHFLIFSSGAGAGSGNQCAVSAA	4030
QY	2358	SNGLTAPNGPSQORVIRQALANAGLSDAVDAEANGSTGTLTGDPIEAQALLATYQDRA	2417	QY	2995	-----	2994
Db	2952	SNGLTAPNGPSQORVIRQALANAGLSDAVDAEANGSTGTLTGDPIEAQALLATYQDRA	3011	Db	4031	NYVLDAATHROTHGLPGASIAMGPWAGKMSAGDAHNGYLEKRGILPMEPRMALAFHR	4090
QY	2418	GGEPWLGVSXVNGHTQAAAGVATKMMALRHGLLPTTLVNDSESPPVDSAGAVOL	2477	QY	2995	-----	2994
Db	3012	GGEPWLGVSXVNGHTQAAAGVATKMMALRHGLLPTTLVNDSESPPVDSAGAVOL	3071	Db	4091	ARAQRNSNLIITADIDMERFVPARTARRSPLEIDIEVYRQAAQELBAASTAKTTTAOR	4150
QY	2478	LTEFTVMPGEGRLRRAGVSSFGVSGTNAHVILEAPADDPVPGPAGSDGASDDEBA	2537	QY	2995	-----	2994
Db	3072	LTEFTVMPGEGRLRRAGVSSFGVSGTNAHVILEAPADDPVPGPAGSDGASDDEBA	3130	Db	4151	IATSLRELRABLTSKQNVLLGLRTGICTVGLNPPGIEBQAFRDLGPDLSIAQF	4210
QY	2538	GSFGVPMVLVSAKSOPRALRAQALNHLNLDHPGLDADVGYTLAARAVFDRATILAA	2597	QY	2995	-----	2994
Db	3131	GSFGVPMVLVSAKSOPRALRAQALNHLNLDHPGLDADVGYTLAARAVFDRATILAA	3190	Db	4211	SKELAKETGLRPLPSLVFDPYPTPOECANHLRQOLVDLDEBDAALSNALPOVARRRVED	4270
QY	2598	DRPTFLOALQALAAGERHRAVHHSARCGTGTGEBAAKTAFICSGGOTQCPGAHGLYHT	2657	QY	2995	-----	2994
Db	3191	DRPTFLOALQALAAGERHRAVHHSARCGTGTGEBAAKTAFICSGGOTQCPGAHGLYHT	3250	Db	4271	EPYALIGMACRPPGVRSADLWELLASGKDAIGVFTDRGMDLTLYDPPDHPGTCYT	4330
QY	2658	HPVFAAALNDICTHLDPHLPLPLTLQ--NNDNEDAAALLQOTFYAOPALFAFOVAL	2715	QY	2995	-----	2994
Db	3251	HPVFAAALNDICTHLDPHLPLPLTLQ--NNDNEDAAALLQOTFYAOPALFAFOVAL	3310	Db	4331	RNGGFLYAGHNPDAEFEGISPREALAMPQORULLETAWETIEHAGINPHTLHGTPGVF	4390
QY	2716	HRLLTQGYHITPHYVAGHSLGEITTAHLAGILTLTDAITLITQATLMQMPGWTTLH	2775	QY	2995	-----	2994
Db	3311	HRLLTQGYHITPHYVAGHSLGEITTAHLAGILTLTDAITLITQATLMQMPGWTTLH	3370	Db	4391	AGINQDHAHNRQSHDVEITBGVALTGSSGSVAGRVAVTGLGEPVAVSDTACSSSLV	4450
QY	2776	TTPHNHTHNLTAENDLAAINTPTSLVSGTPHWQHTTLTLCQOOGITKTKLPTNHAF	2835	QY	2995	-----	2994
Db	3371	TTPHNHTHNLTAENDLAAINTPTSLVSGTPHWQHTTLTLCQOOGITKTKLPTNHAF	3430	Db	4451	ALHMAAOLRAGECSMALAGVTWSSPGTFVEFSSRQGLAADGRKAVSAAADGTGMAB	4510
QY	2836	HSBHTPRLINQLOHQTLYHPRHPTPLTANTPRDOLLTPHYWTOQARWTVUATTTOT	2895	QY	2995	-----	2994
Db	3431	HSBHTPRLINQLOHQTLYHPRHPTPLTANTPRDOLLTPHYWTOQARWTVUATTTOT	3490	Db	4511	GVGMILVERLSDBARRNGHVLAVVRGSAVNDGASNGLTPRNPSSQORVIRQALNAGLT	4570
QY	2896	LHOQGVTTYIELGPDNLTTLTHHNLNPNPTTLTLTPHNNRQTHLLTNLAKTITTWHP	2955	QY	2995	-----	2994
Db	3491	LHOQGVTTYIELGPDNLTTLTHHNLNPNPTTLTLTPHNNRQTHLLTNLAKTITTWHP	3550	Db	4571	PADVDAVEGHGTTLTGDPIEAQALLAAVQGRPHHRPLMLGSLKSNIGHAQAAAGGVV	4630
QY	2956	HHYTHNDNOPHTHNLDTFTYPRFOHNYME-----	2986	QY	2995	-----	2994
Db	3551	HHYTHNDNOPHTHNLDTFTYPRFOHNYME-----	3610	Db	4631	IKVMALRNGLLPQTLHNDPTTPQVDMWSTGAVQLLQTPVPWPADPAGRPHAGVSSFGVS	4690
QY				QY	2995	-----	2994

```

Db 4691 GTNAHILIEBAPFQSDTDDEPPANAPALPHLPVPVSARSEAGLRQAQALRQYVA 4750
QY 2995 ----- 2994
Db 4751 ARPDMSPADIGAGIARGRAVLEHRAVILAADEBELAQALTAAGEBPHITGTRGCD 4810
QY 2995 ----- 2994
Db 4811 RGVVFPVPGQGGQAGMGLTLLTSSPVFAEHIDACKALKTPWVPSLTDLIHRDPDPA 4870
QY 2995 ----- 2994
Db 4871 MQQADVQVPLFSIMVSLAALMSYSGIEPDAVLGHSQGEIAAHICGALSUKDAKTVAL 4930
QY 2995 ----- 2994
Db 4931 RSRALAAVRGRGAMASLPLPAQVQOLISERMEGQMWAAALNGPHSTTVSGDTKAVDEV 4990
QY 2995 ----- 2994
Db 4991 AHCTDTGLRAKRIEVDYASHCPRVQPLHDELLHLSGITPOPSTVPFSTVEGTWLDTTT 5050
QY 2995 ----- 2994
Db 5051 LDAAYWYRNLIHQVPVFSHALIQTLLTDGHRAFIEISPHPTLVPALEDTTENTENITANGS 5110
QY 2995 ----- 2994
Db 5111 LRQDNDTHRELTALAHHTTGTIGTPTTHHNTYQTHPHNPHTHLLPTVPFOHQHWL 5170
QY 2995 -----VSAAAGDPTFHPPLGATLELATDGAALAGRLSIRGHPWLADHAGVGYLIS 3046
Db 5171 QPPTTTDTLTGTLPTHHPLTTLTLLADNNTOLLGRLSIRGHPWLTDHTVAGMVLPR 5230
QY 3047 GATFLELALHAGTVGCDRVDDELTHAPLVVPVGVGVYQVAAADGEGRLVSYARG 3106
Db 5231 GTMLLELALQAGEVDCPRVSELELHAPLVPHIEDVTLOQVTAADSGHRAIAHISYS 5290
QY 3107 GSACGGGAGSGVWTCASGVLYEAAAGGVVVG-----LAGVPPRGANAVIDVGVDR 3161
Db 5291 GTA-----SSADREWTIRHATGTLTHHADTDHRADTHTDACLGSGSWPPPGADPIELGDVYGR 5346
QY 3162 LAG-AGCVLPVPSGSLRAVWMDGDLAAYVCLPREAMGD--AAGFGHLPLALDGVQPLS 3218
Db 5347 MAADSDIAYGPFVFGSLAANRFGDDVLAERVLPEALRDAPAAAFGHPALLDALATA 5406
QY 3219 VLLPGTGFGEGAG-----FGEVVPVAVMGVSLHRAGVTVGRVRSVAVRGGREA 3271
Db 5407 LTPQNGDGTSTENVQESMPDRAAHQARLPFSWGSVSLHTGSSVTLRVRLSRSPQH--NA 5464
QY 3272 VSVVVGEBAGVPVAVSVRLERLPDMQGLRAVSISAGRGSLYAVQMAEV-----GVPV 3326
Db 5465 VALTAABEDGRPVVTLIESLARPSTELRAAADRTREHESLFRLLDWVSVPVANAASPT 5524
QY 3327 CGQAMA-----WHEDEV-----GSGSGGPPVGVVVL-----RCP 3354
Db 5525 ADRWAVIYAGLPHLPGLTEHEHTAYDEPADLLALDRGAPPPGCVLVGGVANTEKREY 5584
QY 3355 DAGAGGGGGGGGGGVEVV-----GGVLGVOVGMGLERFAGSRLVVVTR 3400
Db 5585 SAABRGERGTACARPDVAVGVVHTAAVHAQAQMLARLQAVLGBERLADSLVLTLC 5644
QY 3401 GAVVAGBEDGPDVAVGASVWGLVSAQAEHPRVLLDLDTDTGTDLDTGAGAGWVDGG 3460
Db 5645 GAVVARGSGDADTIDPGAAMVGLVRSASQSEHPRITLLDF--ERSTE-----AEPG 5692
QY 3461 RVAAVVAGCEBQALVBERLLAARLKLSSGDPAPARSGPTRARRSDVPAQBSGVPV- 3519
Db 5693 QLATALNCGERQALVRRGGLFTPLVR-----APRVAA-----VPAVPAVVAPSA 5738
QY 3520 -RRSDVDSGREVLFWLSGGSVLYTGCTGVGAAGVARRHLGAGVGRDILLVSRGPDPA 3578

```

```

Db 5739 GHAAPVPAAG-----PFLPGTVLLTGGTGVGLRVLASHVLEAHGVHRLLLAGRRGPDAEGA 5794
QY 3579 EGRLEALAAAEVRIACQVGBRRRVRLLEGVPACPLTGVVHAAGVLDATIASLTP 3638
Db 5795 PELRAELGGIAGVAVVACCAADROQADLITRIPDRPLTGVHSGHILDDVITSLSP 5854
QY 3639 ERLGVFAAKVDAAALLDELTRGMELSAFVLPSSAAGILGSAGQGVAAANALDALAYR 3698
Db 5855 ERLGAVLRAKADAAALLDELTRGALSFAFMFSASAVVGSPPQGVYAAANAVLDPLAHR 5914
QY 3699 RRAAGLPGVSLANGMEASGWTGHTAGTDHRIITSGHPMSTPDALALFDAAALDRP 3758
Db 5915 RRAEGLPAVSLANGMEBEGGWTGHTDVDHARISAGWRPLPTAALALFDAAALDGE 5974
QY 3759 VLLPADL-----RPAPLPLLODLIPATRRTRRTTGTG--ADNGAQLHARLAG 3807
Db 5975 FLMPARDLTVAVSGAASADVPPLQGLQLPKRSRAAALPGGAPADDAAMWREHLAR 6034
QY 3808 QTHEOQHTTLLALVRSIHATVLTGHTTPTDTPDRARFDLGFDSLTVAVELNRSLRTTGLR 3867
Db 6035 QSAGERRQALLRVRSHVAAVAGHSGADGIDASARARELGFDLTVAVELNRSLTAATGLR 6094
QY 3868 LPTTLAFDHPNPTTLTHLHTLOPOP 3895
Db 6095 LRATLAFDPTPALAEHLGERLLPQE 6122

```

```

RESULT 5
Q9ZG15 PRELIMINARY; PRT; 4613 AA.
ID Q9ZG15;
AC Q9ZG15;
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, last annotation update)
DE Type I polyketide synthase PkAI.
GN PkAI.
OS Streptomyces venezuelae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriales;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=54571;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC15439;
RX MEDLINE=98445333; PubMed=9770448;
RA Xue Y., Zhao L., Liu H.W., Sherman D.H.;
RT "A gene cluster for macrocyclic antibiotic biosynthesis in streptomyces
RT venezuelae: architecture of metabolic diversity."
RL Proc. Natl. Acad. Sci. U.S.A. 95:12111-12116(1998).
DR EMBL; AF079138; AAC69329.1; -.
DR HSSP; P25715; IMLA.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Epanthn-attach.
DR InterPro; IPR000169; SHprot_acsate.
DR Pfam; PF00698; Acyl transferf. 3.
DR Pfam; PF00109; ketoacyl-synt. 3.
DR Pfam; PF02801; ketoacyl-synt C. 3.
DR Pfam; PF00550; pp-binding. 3.
DR PROSITE; PSS0075; ACP_DOMAIN. 3.
DR PROSITE; PSS00606; B KETOACYL SYNTHASE. 2.
DR PROSITE; PSS00012; PHOSPHOANETHEINE. 2.
DR PROSITE; PSS00639; THIOLE PROTEASE_HIS; UNKNOWN_1.
KW phosphopantetheine; Transferase.
SQ SSQENCE 4613 AA; 477281 MW; A3BAREDJ7CEC9383 CRC64;

```

```

Query Match 39.1%; Score 8148; DB 2; Length 4613;
Best Local Similarity 44.8%; Pred. No. 0;
Matches 1909; Conservative 463; Mismatches 1307; Indels 586; Gaps 94;

```

```

QY 15 GEVLGVADADGGVFPVPGQGPQMPQMGRELIDASDVRESVRACEAFAFPVYDMSVEQ 74
Db 620 GLVGVASGV-GVAVFVPGQGTOMAGMGAELIDSSAVFAAABCEAAALSPYDWSLEA 678

```

OY	75	ILRQSPDAPGIDRDVNOVPTLPAWMI SLAALMSQGVPCVILGHSIGEIAAHVSGSL	134
Dp	679	VROQPEGAPTLERUDVNOVPTLPAWMI SLARWQHNGVTPQAVNGHSGEIAAHVAGALS	738
OY	135	LADARVVTLWISOA-QTLTLAGTGA LVSAATPDELLPPIA WTEDNPARLA VAAVNGPRS	193
Dp	739	LDDARVVTLKSISIAHLAOGKGLSLALSBDAVLBRAGDP-----GLSVAAVNGPTA	793
OY	194	TVVSGAREVADVADLTAAOVRTPMIVDVPA NSPLMYAIEBKVSGLBPTTPRPSRIP	253
Dp	794	TVVSGDDPQVQIEELALACACEDGVRARAVIIVDVASHRQVEIIESELAETVLAGISPDARVP	853
OY	254	FHSSVTCGRLDJTRRELDAAVYWRNMSSTVRFERPAR-LLLQGRKTVUMSHPVLTMGLO	312
Dp	854	FFSTLEGWITEPVLVDGGWYRNLRHNGFAPAVETLATDGEFTHFEVSAHPVLTALP	913
OY	313	ELAPDLGPTGTADPVIMGTLRRGCGTLDHFLTSLAQ-----LRGHGET	356
Dp	914	-----GVVTGLA-----TLRNDGGCDRLVASLAEMMAGLAVDMSRPLPSATGHNSD	961
OY	357	SATTVLS-----ARLTALSPT-----QQQSLLLDLVRANTM	387
Dp	962	LPTVAQFQTERHMLGIEIALAPAGEPAVQPAVLRTAEAPRLEDBEQRLVILDKVQAQTA	1021
OY	388	AVLNDGGERNRASDAGSPASFAHIGFDSVMGVELNRLSKATGLRPLPTLLPDTTPPAV	447
Dp	1022	QVL-----GYATGGQIEVDRTRERACSTLTGVDLNNRINAAFGVMA DSMIFDPTPEAL	1077
OY	448	AARLTALAGLDEBTAEPVDPSPSGHGTAAADPRITAIIGNACRPGGVRSRPKYLMELAA	507
Dp	1078	AEQLLVVHNGERAA NPARAEPRPVAAG--AVDEVALVGNACRLPGGVASRPEIDMLRYA	1135
OY	508	SGGDAIGFPPTDRGWPTEQRHAODETOPGTFPYPOGGFPLHDANHDA GFISPREALAM	567
Dp	1136	GGGDAISBFPQDRGMDVEGLYHNPDEHPTGTSYVRQGGFIE NVAGDAAFGISPREALAM	1195
OY	568	DPQQRLLLETSEWEAFERAGIDRPLSRGSRGTGVFAGALSFDYGRPNMDTASSGADAVEGHI	627
Dp	1196	DPQQRLLLETSEWEAVEADGIDPITSLRGQVGFPTGAMTHEBQSL---RDGSGGLDXYL	1251
OY	628	LTGTTGTVSLGRIAVSFGLERPAIVDVGCSASVLTTLHACOSLRSGECTALLAGVSYM	687
Dp	1252	LTCNTASMSGRVSTTGLCEGRBALTVDTACSSLYALHVALOALRKGEVDMALAGVAVM	1311
OY	688	STLGMIFBSRQGLSVDRCKASAAADGTGMBGVMLYERLSDAVRLGHRVLA VR	747
Dp	1312	PTPGMFVFSRQGLAGDGRKAFAPASADGTMSGEGVLLYERLSDBARRNGHQLAVVR	1371
OY	748	GSAYNQDASNGLTAPNPGRPOEVRVTKQALANAGSVADVDVYEGHGTGTTLDGPTLEAOL	807
Dp	1372	GSALNQDASNGLTAPNPGSOQORVRRALADARLTTSDVDVEA HGTGTRLDGPTLEAOL	1431
OY	808	LATYGO-RAGGRPLWLSLKSINIGHTMAAGVGVITKNVMA LREGVTRPTLHVDKPSPOV	866
Dp	1432	IATYGGQHDDEQPLRLGSLKSNIGHTQAAAGVSGVITKNVQMRHGLRPTLHVDEPSQI	1491
OY	867	DMSAGAVVLTLEAVPWPMDAAGRLEARRAGVSSFGIGCTAHYVLEEARAGCSVAGGGVLE	926
Dp	1492	DMSAGAVVLTLEAVDMPEKODGGLRRAVSSFGISCTAHYVLEEARP-----VVE	1542
OY	927	GAPGLAISVAESVAAPVA VASPAVESVFPVVPVVPVVSARSEKGLRQAOEALRQYAV	986
Dp	1543	GA-----SVVPSVYSGS AVGGV-----TPVWVSASAKSAALDAQIERLAFAF-AS	1585
OY	987	RPDVSLAVNGA-----GLACGPAVLEHRVAVVLAAREELVJGIGALMA GEPDRVTT	1038
Dp	1586	KDRITDDAAGAVDAGAVAHVLA DGAQOEHRVAVVLAGAADDVQVLA-----ADBPGL-LR	1639
OY	1039	GHAPGDRGCVFVFPQGGQGMAGKGVLLTASSPVFABRMQACEBALAPWMDVSVLDLR	1098
Dp	1640	GTA SGE--VGRVAFVFPQGGTGMAGKGA LLDSSANFAFAAMAECEBALSSPYVMDSEAVVR	1697
OY	1099	RDAGDAVWERADVOPVLPVWVNSLALMRSTGIEPDAVVLGHSQGEIAAHVCGALS LKD	1158

Dd	1698	QABAPFLERDVQPTTFPAWMISLAWQHNHGTTPRAVGNHSGSELAAAVAGLRLDD	1757
Oy	1159	AAKTVALRSALAA--VGRGGMSVRLPAOEVEOLIGERWAGRLMVAUVNGPRSTAVSGD	1217
Dd	1758	AAAVUTLRKSIANHLAKGGMISLALNEDVALERLSD--FDGLSVAUVNGPRTATVSGD	1815
Oy	1218	AEAVDELVAICAGTGVAAKRIPIUDYASHCHVORLYELLELLDIDROBSGVPRFESTVE	1277
Dd	1816	PVOIEELLAQACKADGFRARIIPUDYASHKROVEIIESELQOVNLGLBROAPRVPFSTLE	1875
Oy	1278	GTWIDTTLDAAYURLHOFVRPSDVAOLLA--DDGRVFEVUSPHYTLVPAIDTDEDT	1336
Dd	1876	GTWITTEVLDSTWYURLNRHVGFAPIETLADDEGFTHFESANHVLTMTLEBT----	1931
Oy	1337	AEDVTAIGSLRGDNDRRLTALANTHTTGIGPTTMNHNHTNNHTPHRTILDRPY	1396
Dd	1932	---VTLGIGTLRREGSGGERLVTLSEAMVWGL--PVAMTSLRLPTABRP-----GLPTY	1980
Oy	1397	PRQONHWLESSORFAGSGSGAGAGSAGSGRAGTAGTAVESRFDVAARODLETVAT	1456
Dd	1981	AFQAEWRULEMTEPRAAALTG-----	1999
Oy	1457	TLAVPRASGLDTVVRALSAHHNQHDARINTWYUOTERWKRYLRTPHORQOT-----WL	1511
Dd	2000	-----DMMKYRIDMK--RLPAEGSERLTSGRWL	2027
Oy	1512	IAIPETOTNNPHNTILNHNHGTPIPLTLNHTTPHONLNTLNTREQOAHNTTG-	1570
Dd	2028	AVTBE--DHSQAAVAVLTALVDAGAKVEVLTAG--ADDRKELAAKRLAL-----TTGD	2077
Oy	1571	AITGLSYLALDETBRNHNHTPTGLLTLTQHTOTPHRRPLWUATTVATTNTPNDR	1630
Dd	2078	GFTGVUSLL-----DGLVRQVAMVQALSDAGIKARLMSVOTQSAVSVGRIDT	2133
Oy	1631	LTHETQOQTGSLATTLLEHRTNAGIIDLRTPTNPTLOHLOTLPORNHQOLARTT	1690
Dd	2124	PADDRLMVLGGLGVVALBENRERWAGVLDLPAORDAALANLTALSAGTEBDIARTT	2193
Oy	1691	GTHTRRLTPTLTPRTTH--OPRTP--TPHGTLLTSGTALATYLNLNHTTPHONLILTS	1747
Dd	2184	GLHARRL---ARARLHRRPRTDMQORNGTULITGCTALSGASHARRM--AHNGAHHLLVS	2239
Oy	1748	RTGERTHBAOHLTQLOOKGILTTTTCTDSNPDOLOQLNLTIRPHNLPTTVITNAGILD	1807
Dd	2240	RSGBQAGCATQLTAELTASGARVTIACDVADPHAMRTLLDAIYAEPTRLGLVHTAGALD	2299
Oy	1808	DATITNLTPROLMNVLAKAHSAHLNQLONHRTPLTFCULYSSAAALFGARGOANYAAAN	1867
Dd	2300	DGIYDVLTAEBVRANRAKAVGASVLDLERTDLDLDFVLFFSVSSTGIRPGOONVAPHN	2359
Oy	1868	AYLDALAHNRHTNLPTATSIAMGTMOGNGLADSPKARAVLDIRGFRMSPRELATAATQA	1927
Dd	2360	AYLDALAAARRATGRSAVSAVAMGWDGGMALDGVALRLNNGVRCMDDELALALABSA	2419
Oy	1928	IADTERPVVALADIDMSK-----IEHTSOTSDLVSA-----ABREPAVORP	1969
Dd	2420	LGRBETH--ITVADIDMORFLIAYSSGRORPLVEBLREVRIIDARSDATSGGSSAQGA	2478
Oy	1970	TPRELHKTALHQTSAQORALLLELVBDHVAVLKRNADPKAIAIDQSGFRALGDSLTAVE	2029
Dd	2479	NPLAE---RLAAAPRGERTTEILGLVBAQAAVLRMSPEVADRAFKDIGFDSLAGE	2535
Oy	2030	FRNLILKATGRLRVLSVPHRPAKALANLONLRG-----TAABASAPSAATAEASV	2084
Dd	2536	LKRNLTATGLOBRATVDFDHRTPLALVSLRSFGLDEBETADARRKRAALRYTGAQAGGA	2595
Oy	2085	-----TEPIAVGASCPREGVTSADDFWDLISSEODAIIGFRTDGRMDLTLTYDRP	2137
Dd	2596	GAGTDDADDPRIAYVAMSCRYGDIIRSPEDLWMMLSEGBEGITPRPTDGRMDLOGLYADR	2655
Oy	2138	DHFGCTTRNGGFLYDAGNDAEFEGISPRBALAMPQORLLLETAMETIEHAGINPHTL	2197

Db 2656 DALGRAYVREGGLHDAEPDAEFFGVSPREALAMDPOORMLTTTSWEAFERAGIEBASL 2715
 QY 2198 HGPRTVFTGNODVALRYVNAAGSTDFALTGTAQSVISGRISYFEGEPRAVSDTA 2257
 Db 2716 RGSSTGVFIGLSTYODYARVNAFRGVEGYLLTGSTBSVASGRATYFGLBEPRTYDTA 2775
 QY 2258 CSSSLVLAHLACALRAGECSMALAGVTWSSPGARVESRQGLADGCKAFSAAD 2317
 Db 2776 CSSSLTLHLAVLRSGECTMALAGVAMATPHMFVEFRQALALPDGSKAFSAD 2835
 QY 2318 GTKGEGVGMILVERLSDAHNRHVRVLAATVGSANODGASNGLTAPNGSQOQVITQAL 2377
 Db 2836 GFGAEGVGLLVRSLDARNGHVRVLAATVGSANGLTAPNGSQOQVITQAL 2895
 QY 2378 ANAGLSADVAVAHGCTGLDPIEAQALLATYGGDRAGEEGLMTGSVSNHGHQA 2437
 Db 2896 ADAALRAGDIDAVATHGTSLSGPTEAQIATYGERPARERLALGVSNSHGHQA 2955
 QY 2438 AGVAVIKMWALRHGLPRTLADDEPSPHYDMGAGAVOLLTETVMPGEGRLRAGVS 2497
 Db 2956 AGAAGIIKMWLVAMHGLPKTLHADDEPSHYDMANSGALVTEPIDMPAGTP-RRANVS 3014
 QY 2498 SFGVSGTNHVLIEAP-----ADVP-----GSPPAGEGDAGSDDEBA 2537
 Db 3015 SFGISGTNAHVLEQAPDAGEVLGADVEPEVSETVAMAGTSGSEVASEAPAP 3074
 QY 2538 GS-----PGWPMVLSAKSOPALRAQAALHANTDHPGLDA-----DYGTLAH 2583
 Db 3075 GSRBASIPGHLPMVLSAKDESLRGQAAALHMLSE-PAADLSADGPARLRDGYTLAT 3133
 QY 2584 ARAVEDHRAVLIADRTFLQALQALAAAGEBPRAVHSSAFGTGTGEAAKTAFIGSQ 2643
 Db 3134 SRFAFARAATVADRDGFLDGLATLAQGGTSAHVHLDIANDGT-----TAFLETCQ 3185
 QY 2644 GTQPGCAAGLYTHPRPAALANDICTHLDPHLDHPLPLTQNDNDNEAALLOOTRY 2703
 Db 3186 GSQPGAGRELVDHPFARALDEICANLDGHELPILVDWFAFEGSAE-ALULDETRY 3243
 QY 2704 AOPALFAFOVALHRLTDGNYHTPHYVAGSHGSEITAAHLAGILTLDATTLTQORNTM 2763
 Db 3244 TQCLFLEVALFELV-BSMKRRPALLSHSGVSEIAAHVAGVSLDAAALVVARGLM 3302
 QY 2764 QTEMP-GTMTTLTTPHITHTLTAHE-----NDLAAIAINTPSTLVISGTPHTVOHITLC 2819
 Db 3303 QELPAGAMLAVOQAEDBIRWLETERVAGRDLVAANVGEAVALSGDAAREARAYM 3362
 QY 2820 QOQCIKTITLPTNHAFFSPHTNPLNQLHQHTOTLTYHPTHTPLTANT-----PPDOLLT 2875
 Db 3363 SGLERRTRALRVSHAFFSAHMDGMLDGFRAVLTEVEFRRPSLTIVVSVTGLAAGPDLCD 3422
 QY 2876 PHWTOQARMTVDYAATTOQLHQGVITYIEELGPDNTLTTLTHNLFPPTTLT---LT 2932
 Db 3423 PEYWRVVRGTVRPLDQVRVLRDVGRTCLELGPDGYLTMAADGLADTPADSAAGSPVG 3482
 QY 2933 HPHHHPOTH-----LLTNL-----AKTTTWHPHYTH-HDNOPHTHT 2969
 Db 3483 SPASPADSAAGALRPRPLVALLRKRSETEYVAALAGRAHAGCTGPDHAMPAGSAGH 3542
 QY 2970 HLDLPTTPQOHNYWLESTOPGAG---NVSAAGLDPTBHPILGATLELADGALLAGRL 3026
 Db 3543 RVDLPYTSFRDRWDA--PAADTAVTAGLIGTDHPLPLGAIVSLPRDGLLTLGRL 3600
 QY 3027 SLRSHPLADHAGVTLGATFLALAHAGTVVGDVDELTLHAPLVVVDGGSVQ 3086
 Db 3601 SLRSHPLADHAGVTLGATFLALAHAGTVVGDVDELTLHAPLVVVDGGSVQ 3660
 QY 3087 VGVAADGE-----GRLVSVYARGSACGGGAGGVWTCASGVLEAAAGGVV 3138
 Db 3661 VTVGAPGEPGEGESAGGARVSLHSLTAD-----PAGTAMSGHAIGLALTDRPELPA 3715
 QY 3139 DGLAGVPPRGAVAVDVGDVDRDLAAGCVLGPVFSGLRAVWDGGLLAECVLP----- 3193
 Db 3716 PDRAAMPPOGAEEVPLDGYERLDGNGLAFGPLFOGILAMVWRVEGEVADIALPATNA 3775

QY 3194 -----EEAWDAG--FGLHALLDGVVCPLEVLPGTGFGGAGGEGVRPAWGV 3246
 Db 3776 TABATANGGSAALAPGIRHALLDASLHAIAN---GGL-----VDEPLVRVPRFHMSV 3827
 QY 3247 SLRAGVTVVRVAVSAGRGGEAVSVVVGDEAGVPVAVDBLERPVDMGRLRAVS 3306
 Db 3828 TVHAAGAARVRLAS---AGTDAVSLSLTDEGEPVLSVERTLRPTVADQAAASRVG 3883
 QY 3307 AGRGSLYANQW-----ABVGPVVGQ-----A 3330
 Db 3884 -----GLMHRVAMPVVALASSGEODPHATSYGPAVLGKBLKVAALLESAGVEGLYPL 3939
 QY 3331 WAMHEDVESGGGPRVGVVVLRCPDAGAGGGGGGGGGEVGVGLGVQGMGLERF 3390
 Db 3940 AALSQV--AAGPARVTLAPL-----AGPADGAGEVGRVATLELLOMLADENH 3392
 QY 3391 AGSRLVVTREGAV-----VAPEDGPVDVVGASVWGLVRSQAQHPDRFVLLDLDTGT- 3445
 Db 3993 AGTRLILVTRGAVRDEPGSGADGDGSHDAAMGLVTRQOTENPGRFGLLDLADASSY 4052
 QY 3446 ---DLDTGAGWGVDCGRVAAVACGEFQLAVRGERLLAALKLESAGVPAQRSGD 3501
 Db 4053 RLPSVSLDAGLR-----DEPOLALHDGTRRLARL----- 4082
 QY 3502 TRARRSPVPAQRSGVPAARSVDVSGREVLPMLSGSVLVTGTVLGAANVHLAGVC 3561
 Db 4083 ---ASVREPTAAR-----LAP---EGVLLTGTGGLGGLVANHVGEMG 4124
 QY 3562 VRDLIVSRGPPAPAGBGLRAELALGAERVIVACDVGERRVRLLEGVPACPLTV 3621
 Db 4125 VRRLIVSRGTPAPADELVHELEALGADVSAACDVADREMLTGLVLDIAPHEPLTAV 4184
 QY 3622 VHAAGVLDATISLPERIGVFAKVDAILDLTLR--GMELSAFVSSAAGILGS 3679
 Db 4185 VHTAGVLDSTLPSMTTEDEVHVRKVDAPFLDLSTPAYDLAFAVWFSSAAVFGG 4244
 QY 3680 AGOGYAAANAAALDALAYRRRAGLPGVSLAMGLMEASGMTGLAGTDHRIIRSGLHP 3739
 Db 4245 AGGAYAAANATDIDALAMRRRAGLPAISGKGLMFTSGMTLOQADLRMRSPACIG 4304
 QY 3740 MSTPDALALFDALALDR-PLVLP-----ADLRPA-----PPLPRLDILLPATRR- 3785
 Db 4305 ISDAEGIALDALRDDRHPVLLPLRLDAAGLDAAGNDPAGI.PALFRDVGARTVAPR 4364
 QY 3786 -----TTRT--TTGANDGQAH-----RLAAGTHEQHTTLALVRSHATVGH 3832
 Db 4365 SAASASTAGTACTPGTADGAETAATVTLADRAATVDGPARRLLEFVVGAEVAGHA 4424
 QY 3833 TPDTIPDDRFRDLGFDLSLTAVALRNLRSRTGLRLPTTLAPDPNPTTLTHLHTOL-- 3890
 Db 4425 RGRIDAREFDLDGDSLTAVALRNLRSAGLALPATLVFHPHSPALASLDELPR 4484
 QY 3891 -QOPDNA-----VAPYLAELDKLESALSALDKTDS-ASERVTLRLK 3930
 Db 4485 GASDQDAGNRNENGTASTARSTAEFDALQLTLREGALVYTGSLDAPGSEVEHLR 4544
 QY 3931 SLMLR-----WNPQRPTESADDD--EKTSATEAIEFK 3963
 Db 4545 SLRSMVTGETGTASGADPGAGSABDRPWAAGDAGGAGGSDGAGVPPPMNASABELFG 4604
 QY 3964 FIDND 3968
 Db 4605 LLDQD 4609

RESULT 6
 Q93H87
 ID Q93H87 PRELIMINARY; PRT; 6048 AA.
 AC Q93H87;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)

DE Modular polyketide synthase.
 GN PTEA1.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; Pubmed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinohe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakai Y., Hattori M.;
 RT "Genomic sequence of an industrial microorganism Streptomyces
 RT avermitilis: Deducing the ability of producing secondary
 RT metabolites."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL: AB070949; BAB69303.1; -.
 DR Interpro: IPR001227; Ac transferase.
 DR Interpro: IPR004410; FAbD.
 DR Interpro: IPR000794; ketocacyl-synth.
 DR Interpro: IPR003880; ppanene attech.
 DR Pfam: PF00698; Acyl transfer. 4.
 DR Pfam: PF00109; ketocacyl-synth. 4.
 DR Pfam: PF02801; ketocacyl-synth. 4.
 DR Pfam: PF00550; pp-binding. 4.
 DR TIGRPFAM: TIGR00128; FAbD. 4.
 DR PROSITE: PS50075; ACP DOMAIN. 4.
 DR PROSITE: PS00012; B KETOACYL SYNTHASE; UNKNOWN. 3.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN. 4.
 KW Phosphopantetheine.
 SQ SEQUENCE 6048 AA; 629420 MW; F86732ED2694705 CRC64;
 Query Match 37.9%; Score 7891; DB 2; Length 6048;
 Best Local Similarity 45.7%; Pred. No. 0;
 Matches 1838; Conservative 443; Mismatches 1269; Indels 474; Gaps 83;
 QY 26 GGVVFPFGGQPGWPGKRELLDASDVRESVACBAFAFYDWSVEQVL---RSPDA 82
 DB 548 GELAFLEPGGSGQSGAKGRELASVPAFASALDVCNLDPLPRPLREVWFAVVGGEA 607
 QY 83 PGIDRVUVOPTLPAVWISLALMRSGVPCAVLGHSLGEIAAAHVSGLSLADARV 142
 DB 608 AELDQTLYTQTSLPFAVEVALFRLESWDITPDVLMGSHIGELAAAHVAGVSLADACALV 667
 QY 143 TLMSAQTTLACGALVSVAATPELLPRIAPWTEDPARLAVAAVNGPRSTVSGREA 202
 DB 668 AARGLMQALPAGGAMIAVDADEVARIGERTD---RVSIADLGPSSVVSSED 723
 QY 203 VADLVADLTAAOYTRMI PVDVPAHSP---LWYAIERVVSGLPITPRPSRI PFHSSVT 259
 DB 724 ATFIADAFALGRTSRLRVSHAHSPHMERMLAEFNVAGL---TFHAPRIPIVSNVT 780
 QY 260 GGRLDTRE-----LDAAYTRNNSSTVRFEPARLLIQGPKTFVENSPPHVLTMGLQEL 314
 DB 781 -GRLSERPSGEGCTAAVWVRHVEAARFADGVARLDEOGVRYTLELPGGVLA----- 833
 QY 315 APRLDGTTGADTYIMGLTRGCGTLDHPLTSLAQLGHG----- 354
 DB 834 --SMAAGADSESLFVPAKTRRPEAOALLAAVASLHVHGVDPMDALFGCGGARRKVA 891
 QY 355 -----ETSATTV--LSARLTALSPTOQ 375
 DB 892 LPTPAFERORVWLDGTTSPVATTGMARVAPORPAIETDSAPRGALGRLLAGLAESE 951
 QY 376 SLILDLVRAITMAVLNDGNERITASDAGPSAPHLGPDVSMGVELNRLSKATGLRLPV 435
 DB 952 EAVLDVWRAHIAVTL-----GHAEARQVETETPFKDLGFDLSISVELNOLSEATGLRPS 1007
 QY 436 TLFDHTTPAAVAAARLTALGHDEDTAPVPSDPSGHGGTAA-----DDPIAII 466
 DB 1008 GLRFDHTPAAHLAHLAOSIG-----GGDAAAARATYAADHPERPAIV 1051
 QY 487 GMACRPFPGVRSRKLWELAAAGDAIGPRTDGRMTEQRHADPTQGTFFYFGGSGFL 546

DB 1052 AMSCRPLPGVSSPEDELMRLASGDSIGSPDTDRGMDVETLYDPDEGTPEKSTRHGFL 1111
 QY 547 HDAHNDAGFFGISPREALMDPOORLLLETSEAEFERAGIDPLSRGSGTGFAGLSLF 606
 DB 1112 HEAAEFDPAGFPGISPREAAIDPQORLVLETAWEAEFERAGIDPATLRGSGAGVFIGTAAQ 1171
 QY 607 DYGRMDTASSSEGAAVEGHILGTGTVLSGRVASFGLGEPATVDTGCSASLVTLHL 666
 DB 1172 DYGRPL-----HEPADVEGILGTGTVSASGVASFGLEGAVVDTACSSVLAHL 1227
 QY 667 ACOSLRSGECTLALAGVSVMTLGMFIEPSRQGLSVDRCAVYAAADGTWGCBCVM 726
 DB 1228 AVQSLRSGECTMALAGVVTMSTPGMFVFSRQGLSADGRCAFAAGADGTGMAEGVM 1287
 QY 727 LVERISDAVRIGHRIAVVRGSANVODASNLTLPNPAGORRVROLANGLSADV 786
 DB 1288 LVERISDARRNHOVLAVVRGSANVODASNLTLPNPAGORRVROLANGLSADV 1347
 QY 787 DVEGHGTGTLGDPTEAOALLATYGC--RAGDRPLMGLSKNIGHTMAAGVGVYKVM 845
 DB 1348 DAVEAGTGTLLGDPTEAEALIAVYQGRDGRPLGLSKNIGHTMAAGVGVYKVM 1407
 QY 846 MALREGVLPRTLVHDXSPQVDMSAGAVLLTEAVPWPGDAGRRLRAGVSFGIGTNA 905
 DB 1408 MAMRHGVLPRTLVHDXPTPHVDMSAGVRLTEAVEMP--ARDLPRAAVSFGIGTNA 1465
 QY 906 HVLEEA-----PAAGCVAGGVLEGAQGLAISVESAAPAVASAPVAPVPPVP 960
 DB 1466 HVLEEAATSEERAKRTAAV-----P 1487
 QY 961 VPVVARSASEAGRAQAEALROYV-----AVRPDVSADVAGIACRAVLERRAV 1012
 DB 1488 VPMVVAKSDVALREDAQRLSLFLDAGVCAQVRB---DVALLATSRALERRAAV 1543
 QY 1013 LAADREELVQGLALAGEPDRRTTGHP--GGDRG--VFVFPGGCGGAGKGVRLAS 1070
 DB 1544 VGGDLAEIKQGLEALASG-----ASGAAPVGGAVGGRVGFPLSGGSGQRIKMGRELIAA 1597
 QY 1071 SPVFARMQACEALAPWYDMSVVDILRRDAGAVVERADVOPVLFVWVSLALMRSY 1130
 DB 1598 YPVFA---AAVDECAFIDAPV-----DVSEELNQTGTPALFAVEVALFRLESW 1647
 QY 1131 GIEPDVVLGHSOEIAAAHVCALSLDKAAKTVALRSRLAAVRRGGMASVLPDAQEVE 1190
 DB 1648 GIEPDVVAHGSVEIAAAHVAAGVLSLEDAKLVSAALMQALPDGAMAV---QATE 1703
 QY 1191 QLIGERWAGRLWAAVNGPRSTAVSGDAEVDVLAACAGTVRARIRIPDVASHCPHYO 1250
 DB 1704 DEVLRHLTDGVGIAAANGPRSVVVSAGABDVTAVGEI PAGRGRKTSRLKVSHPFLMD 1763
 QY 1251 PLREBELLELGDSPSPSGVFPFSTVEGTMLOTTTIDAAYWYNNLQVRFSAVQALAD 1310
 DB 1764 PMLDEFAAVVVRGLTFSESRI PIVVSNLTGRLAETVTE--YWNHVEAVFAAGIEFLGE 1821
 QY 1311 DGHVFEVVSPPHVLPAIEDTTEDETAEDVTAIGSLRGDNDRRPFLTALAHTTGIGT 1370
 DB 1822 LGVTFVEVGPGLV--SALAQSGFDGTDITVYVPLRADRREYALVTALGRHTHGIS- 1879
 QY 1371 PTTMHHYTHNHTPHRPHNTLDTLPYRFQOHYMLSSOPGAGSGGAGAGSGAGS 1430
 DB 1880 -PDMQVLPFGAR-----RVDLPTVAFQERWMLAP-----RAA 1912
 QY 1431 TAGGTAEVSRFPDVAVADOLEVATTLAVPBAGLDVTVPVLSAHHNHQDARINTYT 1490
 DB 1913 TAGAPDMAAEFWDVSESEBRASLIGLLEPTE--LDVVASKLSAARRKREOSTTDGMR 1971
 QY 1491 YQETMKRPLPTTHOHOQWTLAIPTQTHNPHITLITLVNHNHGTPTPLNHTHTNP 1550
 DB 1972 YRITMQPLDDMVAASGTLVYVSEETAETAIKAGLTEL---GVALVFAVGBOTDR 2028
 QY 1551 OHLNHTLHNTROQONHTTGATLGLSLALDETPRPHNHTPTGLNLTLTQHTQTH 1610

Db 2029 ALAFAADAGOEQADGILLFAAAGEAGANALORL-----VLVQALDADG 2073
 QY 1611 PPRFLWATTNATTHNDPLTHPTOAGTWGLATTILLEHPTTAGIIDLPTTPTHTIQ 1670
 Db 2074 IEADLWCLTSGASTGADLTDPVSAQWGLGRIVALBOFORMGGLVDPAEIDSRAL 2133
 QY 1671 HLTOQLTPHNOTOLARTTGTTHRL--TPTTLPTHOPPTPYHGTTLITGSGALAT 1728
 Db 2134 RLGLGLAQ--DEQDLAVRASGVRLVRAPLSAAPADAP--WSPRGVLVTGSGALG 2190
 QY 1729 HLTHLTHPTOQLHLTSTRGPTPHAOHLTTLOQKGLHLTTTCDSNPDOQLN 1788
 Db 2191 QVARNL--AGAGEHLVLTSRGPDAPGAELKAELEIAGQVTVVACDADRDRLALF 2249
 QY 1789 TIRPHLTVTHAGLIDATLFTNLPTOLANVLRKASAHLLHOLTQ--HPTPLFVL 1847
 Db 2250 ----RIVNAVHTRAGVLDDELVESTLPERLNHVRPVDALHETRLRODLDFVL 2305
 QY 1848 YSSAATFGAPGQANVANAAYLDALAHHRHTHLPATSIAMGTWQNGLADSDKARAYL 1907
 Db 2306 FSSMTGVWNGGQAGAYGAANFLDALAEHRAQGVPLAVALWMSMGADGMDG--AAGDHL 2364
 QY 1908 DRGFRMSPBELATAVTAOTAIADTERPVYVADIWSKIHT---SQTSL--VSARE 1961
 Db 2365 RRRGVRAIALPALISVHGHGALTHEGETS--MTVADVMEFVPAFVGTRPCPLLGVPEARE 2423
 QY 1962 ----REAPVORP--TPRAELHKTLLAHQTSADORALLLEVRDHVAVLRHADPKAIADDS 2016
 Db 2424 ALERARPAOSPEVVSALVORLGAAPREGGRILLDLVREQAATVGHGKGFBEADRT 2483
 QY 2017 FRALGFDSTLAVEFRNLILKATGLRLVSLVFDHPTPAKLAHLONOLRG--AAESAPSA 2075
 Db 2484 FREGFDSLAVELRHRNLNATGLRLPTLVFHDPTPALARHGEELHGHGETA--A 2541
 QY 2076 AATYAKSVTBPRIAYMGACFPFGCVTSADFMWLISSEODATIGCFPTDRGMDLTYDP 2135
 Db 2542 AQATPAALAADEBPLAIVKMSCLPEGVOSPEDLMELLSGRDAMSGFPTDGMIDLAYDP 2601
 QY 2136 DPHPGTCYTRNGGFLYDAGHFDAPFPGISPREALANDPOORLLETFAMETIEHAGINPH 2195
 Db 2602 DDPFRGKTSREGGFIQADRFDAGLFGISPREALANDPOORLLETFAMERFEPAGIALA 2661
 QY 2196 TLHGPTFTGTNGODYALRVHNAQSTDGFAITAGASVTSGRISYTFGFBGPAVSVD 2255
 Db 2662 SVRASRGVFLGTNGODYANGLRNAPREIEGYALTGAASVSGRISYTFGFBGPAVSVD 2721
 QY 2256 TACSSSLVALHLACOALRAGECSMALAGCTVWSSPGAFFEFSPORGLAADGHCKAFSA 2315
 Db 2722 TACSSSLVALHLAOLRSGCSMALVGVYVMTTPTDLFVEFSPORGLSPDGRCKAFSA 2781
 QY 2316 ADGTGMEGEGVMLVERLSDHNRGHVLAVVREGSANODGASNGLTPAPGSPQORVIRQ 2375
 Db 2782 ADGTGMEGEGVMLVERLSDHNRGHVLAVVREGSANODGASNGLTPAPGSPQORVIRQ 2841
 QY 2376 ALANAGLSADVDVAEHAHGTCTTGLDPIEAQALLATYAGODRAGEGRLMGSVKSNGHTQ 2435
 Db 2842 ALASAGLTPADVDAVEAHGTGTRLDPIEAQALLATYAGOHSDRPAWLSVKSNGHTQ 2901
 QY 2436 AAAGVAGVITKVMALRHGLLPRTHVDEPBPHVMSAGAVOLLETVPWPGBSRLRAG 2495
 Db 2902 AAAGVAGVITKVMALRHGLVPLTHVDEPBPHVMSAGAVALLLEPQWGGTGP--RRAA 2960
 QY 2496 VSSRGVGTNAHVLEEARPADDVPGGPRPAGBGDAGSDEAAAGSPGVMPVLVASKSPAL 2555
 Db 2961 VSSRGVGTNAHVLEEARPADDVPGGPRPAGBGDAGSDEAAAGSPGVMPVLVASKSPAL 3007
 QY 2556 RAQOALHAHLTHPGLDLDVGYTTLAARAVFDPHRTLLAADRTFLQALQALAAEPH 2615
 Db 3008 RAQOERLRAFTD--DVSVDAGGSSATTSALHRAAVVIGSDAEIRAGIEALAAEPH 3065
 QY 2616 PAVIHSSAPGGTGTGEAAGKTAFLCSGQGTORPGMAHGLYTHPVPFAALINDICTHLDPH 2675
 Db 3066 ANVV-----AGRAHSADKVGFLFSGGSGORLGMRELYAAAYVFAAADVCAALIDAT 3118

QY 2676 LDRPLLPLTQNDNEDAAALLQOTRYAPALFAFOVALHRLITDGYHITPPYVAGSHL 2735
 Db 3119 VVDV-----AETHOTGSTORALFANVVALFRL--ESWGRIPDYVAGSHV 3162
 QY 2736 GEITAAHLAHLITLTDATTLITQATLMQTPR--GWTTLHTTHPHHITHLTAHENDLAI 2794
 Db 3163 GEITAAHLAHLITLTDATTLITQATLMQTPR--GWTTLHTTHPHHITHLTAHENDLAI 3219
 QY 2795 AATNPSLVISGTPHTVQHITTLCOOQGIKKTITLPTNHAHSPHNPILNOLHQTQL 2854
 Db 3220 AATNPSLVISGTPHTVQHITTLCOOQGIKKTITLPTNHAHSPHNPILNOLHQTQL 3279
 QY 2855 TYHPHTPLTANTPP--DQLTPTHYTQOARNTVDATTTQTLHOHGVTYIAGDNTL 2913
 Db 3280 TPEPMTIPVSNLTGLTBSYTBGVYVRAHREAVRADGVTLHELGVTTFVIGCGCVL 3339
 QY 2914 TTLTHNLBN-----PPTTLTLTHPHHNPOTHLITLAKTTTWHPHNYT 2959
 Db 3340 SALVQGLDGDGTIVITIPALRADRPQALTTAYAOQH-----VSGAEVWMAFF-- 3389
 QY 2960 HHNOHHTHLDLPTTPQHNNHYLSTOPAGNVSAGLDPHTERPLGATLELATDGG 3019
 Db 3390 ----FGAR--RVBLPTFAFORERYWDA--PASAGDMRAVQOGBAGHPLGAAPLADSVG 3442
 QY 3020 ALLAGRLSRHPMLADHNAVGTVLISGATFLELALHAAGVYVGCDDVDELTHAPLVV 3079
 Db 3443 HLITGRLSKTHMLVNDHVASVALLPRTAFVBLATTAADAVCDLLEILITLTPLVPE 3502
 QY 3080 DGGVSVQVGAADAGGRRLVSVYAR--GGSAGCGGAGSGWTCASGVLYEAAAAGV 3137
 Db 3503 RGVVALQVHVGADDGGRRLVTHSQIIDGDAAHNBDEBRANVRAASGVLTVAATG--P 3560
 QY 3138 VDLGAGVPRGVAIVDVQVPRDLGAGCVNPGVSSGRLVAYRDCGDLLEFVCLPEAN 3197
 Db 3561 VDSLTPAMPVGAEPVDVGFYRLAGMALDYGPVQGLRAARSDDPFPAEVLPGAER 3620
 QY 3198 GDAAGFHPALD-----GVQPLSVLEGGTGFEGGAGFEGGVPRVAVMGVSLH 3249
 Db 3621 TDAGATGLHPLDALHLYWGLAVER-----EAGTGNGL--LPRFMSGVRLA 3666
 QY 3250 RAGVTGRVRSVAVGCGGGEAVSVVVGDEAGVVASVDRLELRYVDMGOLRAVSAGR 3309
 Db 3667 AAGASDLRXYVSRAG--GHTVSLVLDGTGPAVEVSLTLRPPADQLRA--GSGS 3721
 QY 3310 RGSLYAVQMAEYV--PVPVCGQAMAHEDVY-----BEGGPRVPGVVTLRCPDAG 3357
 Db 3722 DDVAFGLKMTPVSLPSPAPGMRLETYEDLAALRATDPADTEADAAAADAVVPCPT-- 3779
 QY 3358 AGGGGGGGGGGGEVVGVLGVVQGMGLGERPAGSRLVVVTGAVVAGBEDGPVNVGA 3417
 Db 3780 ----GTADLATRREYTHAVLELVQWMLAEER--PGRVLVTRG-----DLAOS 3824
 QY 3418 SYMGLVRSQAQEHDPFVLLDLDTDGTDLDTGAGAGWVGGRVAAV--VACGEPOL 3473
 Db 3825 AVCGLRSQOTENPDRILVYE-----TSSDAHAGSGVDADLARVLPGLIASGEPQ 3878
 QY 3474 AYRG--ERLLAARUKLESSGDVPAORSCTRARSVPPAORSRGVARSVDVSGREYV 3532
 Db 3879 AYHADEVRVRLRV-----AVP-----DABEP 3904
 QY 3533 WLSGGSVLVYTGTVGAAVARHLAGVGRDILLVSRGPDAPGAGEGRAELAAAGV 3592
 Db 3905 GLDSGTVLITGASGGIGLPARHLVAHEGVRSLILVSRGPDAPGAGEGRAELAAAGV 3644
 QY 3593 RIVACDVERREVRRLLEGPACPLTGVVHAAGVLDATITASLTERLGTVPAAKVDA 3652
 Db 3965 TMAACVADRAAVALLAG--PGQSLAIHTAGVLDGIGSLTERLDAVRRPKVDA 4022
 QY 3653 LILDELTRGM--ELSAFVLPSSAAGILGSAQGNVAAAALDALLAYRRRAAGLPGVSLA 3710
 Db 4023 LNLHDLAGEISGLLSAFVLFSSVAGTIGTPGQANVAAANTFLDALAEHRRAGORATSLA 4082

QY 3711 WGLW--EASGWTGHLGADTHRRIRISGLHPSMSTPDALAFDALALDRPVLLPADU--- 3765
DB 4083 WGLWAGSBSAMQSLDHTDLTRIQRMGLAIPSSDGLRFLFDALALDRVAVLRLDLG 4142
QY 3766 ---RPAAPLPPLLODILLPATRRRTTTRTTTGGADNGAQLHARLAGTCHQOHTTLALVR 3822
DB 4143 GLOCTVPVLRGRLGTGTARARRAQTAPKG-----SLGQRLLGLABAREQVVDLVR 4196
QY 3823 SHIATVGHHTTPTDPTIPDPAFRDGFPSDLTAVELRNRLSRTTGLRLPTTLAFDHPNPTTL 3882
DB 4197 TEVAAYVGHAGAGSVGADDSKDIGFDSLTRAVELRNRLNSAVGKRLPATILLFDYPNPTL 4256
QY 3883 THHL 3886
DB 4257 AHFL 4260
RESULT 7
ID 033954 PRELIMINARY; PRT; 4472 AA.
AC 033954;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DE TyLactone synthase starter module and modules 1 & 2.
GN tyLG.
OS Streptomyces fradiae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1906;
RN [1]
RP SEQUENCE FROM N.A.
RA DeHoff B.6., Sutton K.L., Roestek P.R. Jr.;
RT "Sequence of Streptomyces fradiae tyLactone synthase gene tyLG."
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78289; AAB65504.1; -
DR InterPro; IPR002106; AALRNA_11gaee11.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR002198; ADH short.
DR InterPro; IPR000794; ketoacyl-synt.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR003880; Prantne_attach.
DR Pfam; PF00106; Acl_crenaf; 3.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt; 3.
DR Pfam; PF00550; bp-binding; 3.
DR PROSITE; PS00339; AA_TRNA_LIGASE_I1_2; UNKNOWN_3.
DR PROSITE; PS00075; ACP_DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 2.
KW Phosphopantetheine; Transferase.
SQ SSQUENCE 4472 AA; 463275 MW; E911D1B9E0C332CF CRC64;
Query Match 37.0%; Score 7705.5; DB 2; Length 4472;
Best Local Similarity 44.3%; Pred. No. 0; Indels 495; Gaps 97;
Matches 1849; Conservative 453; Mismatches 1380;

DB 743 ASTVSGDRRAVAGVAVCAQGVQARLI.PVDYASHSRHVEDLKGELERLVSG---IRPR 799
QY 249 PSRIPIHSSVATGRLDTRBELDAUYWRNMSSTYRFPBARALLIQGQPKFVMSPPHVL 308
DB 800 SPFRVPCSTVAGQPGEPVFDAGYWRNLRNVEFAVVGGLLEBGRRIEVSAAHVLV 859
QY 309 MGLQELAPDLGDTTGTADTVI--MGLYRRQGLDHFRLTSLAQRGHET----- 356
DB 860 HALEQ-----TAAADRSVHATGTLRRQDSPHRLTSTABAMAHGATLWDPALPG 912
QY 357 SATTV-----LSARLTLSPTQOO-SLLDLVRA 384
DB 913 HLTLTPYPPNNHNYLDTIDGGGDATQEKESGELTRELRLPSOKQLGLDLVCR 972
QY 385 HTMAVINDQGNERTASDAGPSAFHILGFDPSVWGVLELRNLSKATGRLPVTLIPDHTP 444
DB 973 HTAVVGLD---TAEVDPDLSEFKQGIQSMGVELRNRLTETGLALFTTLVYDRPT 1028
QY 445 AAVARLRTAALGHLDEDTAPVDSPSGHGTAA---ADDPATIGNACRPPGVRS 499
DB 1029 RALAQLHTELL-----DGSPSGVLAFAQKSFQAQELIYAVGMCGRPGVGS 1078
QY 500 KDLWEI.ASGGDAIGFPPTDRGWPTQRHAQDPTQGTYPQGGFLHDAHFDAGFC 559
DB 1079 EALMRVVEGVDAVSEFPDGRGWDVGLYDPBEGVAGKSYVREGFLHDAERDAEF 1138
QY 560 SPREALAMPQORLLLETSWEAFERAGIDPLSVRSGRSTGFAGALSPDVGPRMDTAS 619
DB 1139 SPREAVAMDQQLLETSWEALERAGIDPHSLHSGRTGYAVAMPQOEYPRL---AEG 1194
QY 620 AADVEGHILGTGVSGLSRIAYSFGLGPAITVDTCSSASLTTLHLACSLRGECTLA 679
DB 1195 AESDSQYLLTGISGSVSGVATTLLEGAIVYDTRACSSSLVAHLAVALGAGBEDMA 1254
QY 680 LAGGVSMSTLGFIFPSRQGLSDVGRCKAASAAADGTGSGVGMVLVERLSDAVRLG 739
DB 1255 LAGGVYTMAGPGMFVFFSRQGLAADRCFAFDAGDGTMAAGAGVVLVERLSDA 1314
QY 740 HRYLAVRGSVAVNODASNGLTAPNGPAQORVROLANAGLSVADVVEGHTGTTG 799
DB 1315 HPLAVAVCGSAVAVODASNGLTAPSGSQQRVROLANRLVLADVDDVEAHTGTGRLG 1374
QY 800 DPTEAQLATATYQC-RAGDRPLWLSLKSNIQHTMAAGVGYIKWMLAREGLVPTLH 858
DB 1375 DPTEAQLATCTYGRDDGGRPWLSLKSNIQHTMAAGVGYIKWMLAREGLVPTLH 1334
QY 859 VDKPSQVDMASAGAVRLTEAVPWPGDAGRLRAGVSSFGICGTNAHVILEBAPAA 918
DB 1435 VDEPSRHVDMSAGVWLTEAREWPG--VDRPRAAVSAFVGSGTNAHLILE----- 1484
QY 919 VAGGVLEAGRGALISVAESVAPVAAPVAESVPV--PVPPVPVPVARSBAQLRAQ 976
DB 1485 -----APDTAEASATTPVRSEVSSAAVLDARSQVPPVVGSRVVVREA 1531
QY 977 AEARQVAVRPVDSADVAGAGIACRAVLEHRAVLAADREBLVQIGALAAEGPRRV 1036
DB 1532 AGRLAEVNER-GGVGLADVAVTMA--GRSRGYAAVVLAREBALAGRLRLAGDDPAGV 1589
QY 1037 TTGHA---PGDRGGVVPVFPQGGQWAGVRLLASPFVFAARRQACEBALAPVVD 1092
DB 1590 VTGAVVDPETSGGGGVVLVFPQGTQWVMGAGLLGSSSEVFAASRECARALSVHEMD 1649
QY 1093 VVDILRRDGDADWBERADVQPLPFSVMVSLALMSYGTBPAVYVCHSGGCTAAAHV 1152
DB 1650 LLEV---SGAGLEVEDVQPTVMVMSLAVYKQAMGADVAAVVGHSGETIAAATVAG 1706
QY 1153 ALSLKDAKTAVALRSBALAA--VRGRGMAVPLPAQVEVLEIGERNAGRLMVAANVGR 1211
DB 1707 ALSLEDAANVALARAGLIGRYLAGRGMAAVLPAPAEVEVGLA-KRPG-VBVAANVGR 1764
QY 1212 TAVSGDAEAVDEVLAACAGTVRARIPVDYASHCPVQPLRELLLELGDISPQSGVP 1271
DB 1765 TVVSGDRRAVAGVAVCAQGVQARLI.PVDYASHSRHVEDLKGELERLVSGIRPRSPRV 1824

QY 1272 FFSTVEGTWLDTTTLDAAYWYRNLIHQPVRFSDAVOALADGHRFVEVSPHPTLVPAIED 1331
 DB 1825 VCSIVABEQBGEPEFVFDAGYWRNLRNVERFSAVVGLEBGRHFIEVSAHPVLVHAEQ 1884
 QY 1332 TTEDTADVTALISGLRKGNDTRRPLTALAHHTTGTIGPTTTHHHTTHHTHHPHTHL 1391
 DB 1885 TAAEADSVAHATGLRRQDSDPHRLITS-----TAAHMAHGAITLWDPALPGHL 1934
 QY 1392 -DLTYFPOHNYWLESQPGASGSGAGSAGSAGAGTAGTAEVESFMAVAVAROD 1450
 DB 1935 TLTFTYFNHHWLDTP----- 1953
 QY 1451 LEVATTLAVPSPASGLDTPVVALSAMRHQDQARINTWYQETMKPLTPTTHQPHQ-- 1508
 DB 1954 -TTPATTTQSPDA-----WRYVVMKALITSSPVPRPSIG 1988
 QY 1509 TWLIAPEOTQHDPHTNLTNLHHNGIT---PIPLTNTHTNPOHLNHTLHHTROQA 1564
 DB 1989 RCLLVAPPT-TDGEILDGLTTLSEBGASVAREVPIGARPA-----EVALLKPSMESA 2042
 QY 1565 QNHTGATITGLSLADETPHHPHTPTGTNLNLTLOTHQTH--PPTPLWATTNAT 1623
 DB 2043 GEENT-TVSLGLV-----PST-DAVTSIALQVSDIGVAPARVWALTERRAV 2090
 QY 1624 TTHNDPLTPTQATGLARTTLLEHTHTAGIIDP-----TPTPTLLOHLT 1673
 DB 2091 AVVGEF-TPODAGQWNGFVVALBELPDIWGGIIDEPELAEITRTETSPPTPRRLP 2149
 QY 1674 QT-----LTPRHQTOAIRTTGTHTRRLTPTTLPTTHQPTPTPHGTTLITG 1721
 DB 2150 QTPRRALTELAVALAGRDGEDQVAVRASGIYGRVRAAAGA---ASWQSGTVALTG 2206
 QY 1722 GTGALHTLHTLHTHOTOHLTLTSRGTPTBHAQHLTTOLOKHILITTTCDTSNDP 1781
 DB 2207 GMGAIIGRLARLAA-EGAEKLVLSRGPAPAAELAEELRKHGCEVHAACDVARD 2265
 QY 1782 QLOOLNTIPROHDLTVIHTAGIILDDATLTNLTPTOLNVLAKAASAHILHTOTHP 1841
 DB 2266 ALALVTRAPP---NAVHTAGILDDAVIDTISPESEFVYRGAKVCGAELHOLTDADIK 2321
 QY 1842 -LTAFLVYSSAAATFGAPGQANVAANAYDALAHHRHTHLPATSIAMGTWQNGIADS 1900
 DB 2322 GLDFVFLFSSVGTWGNAGGAYAAANALDALERRBAAGLPATSVAMGIMGGGMA-A 2380
 QY 1901 DKAAVYDRGFRMSPELTAATQAIADTERPVYVIADIDMSKIE-----HTSQSD 1954
 DB 2381 GAGEESLSRRGLRAMDPDAVDALLGAMGRND-VCTVVDVDERFAPATNALIRGRLPD 2439
 QY 1955 LVSAAREREPAN-----QRPTRAEIHKTLAHQTSADQRAALLLEVRDVAALRHADPKA 2010
 DB 2440 TVPARAEALTAGTTSATPDQAPRLARLSMLNTEBRLKVELVREBAFVLRHPMTDA 2499
 QY 2011 IAPQSFALGFDSLAVEFRNLIIKATGLRPLVSLVDPHTPAKLAVHLQNUOGR-- 2067
 DB 2500 IGARPEKSGFDDLTLELNRNLNAGTGLPATVIFDHPSPALARILLDLRLTGACAP 2559
 QY 2068 --AAESAASAAYTAESAVEREIAIVGMACRFPGGVTSADDFMLLISSEQATIGFPTDR 2125
 DB 2560 APADDEPLVAVAD---DPVAVIGACRFPGAGCPALMWLVTVEERVIIGAAPTDR 2615
 QY 2126 GMDLDYVDPDPDPTCYTNNGFLYDAGHFDAEFGISPREALAMPPOORLLLETFAME 2185
 DB 2616 GMDLDYVDPDPGAGTYYREGGFLHDAEFDAEFGCSIPREAVANDPOORLLLETSWE 2615
 QY 2186 TIEHAGINPTLHGTPTGVFTGTNGQDYALRVNAGOSTDGFALTGTAGSVISGRISYTF 2245
 DB 2676 AIEBAGIDPHSHLSRGTGVVGLTHQEYASRLHEAPEYEGYLLTGKSASVSGRISYTL 2735
 QY 2246 GFEBPAYSVDTPACSSILVALHLACOLRAGECSMALAGVTWMSPCAUFERSQRIAA 2305
 DB 2736 GLBEPSSISIDTACSSILVALHNAQAOLRGGCDMALAGVTWMAAPGLFVEFBSRQRIAA 2795

QY 2306 DGHCKAFSAADGTGNGEGVMILLVERLSDAHNRHVRVLAVYRGSANVODGASNGLTAPN 2365
 DB 2796 DGRCKAPADADGTAMMEGAGVULVERLSDBARLGHVLAIVVCGSAVNOGASNGLTAPS 2855
 QY 2366 GPGQOYIRQALNAGISADVDVAVEAGCTTLGDPTEQALLATTGQORABEGPLWLG 2425
 DB 2856 GPGQERYIRQALNARLTVAADVVEAGHGTGLSGPIEQALLGYGRDRDACEPWLIG 2915
 QY 2426 SVKSNVHTQAAAGVAVGIVGMALRHGILLPRTLHDEPSPHYDMSAGAVOLLTEVPWP 2485
 DB 2916 SLKSNIGHQAAGVAVGIVGMALRMKGYMLPRTLHDEPSRHNDMSAGVRLITEAREWP 2975
 QY 2486 GGEGRRLRACVSSFGVGTNAVHLEBP-----ADVPGPPPAGE----- 2527
 DB 2976 GVD-RPRRAVSAFVGSTNAHLIL-EAPALEALEXTADPEAPEAREADVTDTALE 3033
 QY 2528 --DAGSDEA-AAGSP-----GWPMLVSASQPLRQAOALNHLTDHGLDADV 2577
 DB 3034 APDATEBEGAKAPSEBQPAVGVPVNVSGSRVVRBAAGRL-AEVEAGGVGLADV 3092
 QY 2578 GYTLARAVEDHRAITLADRDTPLQALQALAGEPHPAVHSSAPGCTGEAAGKTA 2637
 DB 3093 AVTMA-GSRFGRVAVILANGAEALGRRLAGGPDAGV-TGAVVDETSGGGGV 3150
 QY 2638 FICSGQSTQRPMAHGLYHTHPVPAALNDICTHLDPHLPLPLTONDNDEDAAL 2697
 DB 3151 LVFPQGTQWVGAGILGSSVEFPAASMRCAVALSVHEWDLLEV-----CGAG 3202
 QY 2698 LQOTRYAOPALFAPVALNHLTDGHNITHTHYAGSLGSIITNAHLAGITLTDTATLIT 2757
 DB 3203 LERDVQOPRTWAVMSLAR-YWQAMGVDAVAVGHSQGIATAATYAGALSLDDAALVA 3261
 QY 2758 ORATIM-QTMRPGMTTLTHTPHHTHNLTAHENDALAIATPSTVSGTPTVQHI 2815
 DB 3262 LRAGLGRYLAGGMAAAYELRAGEVBAGI-AKRPVQVAVAVNPGSTVSGRRAVAGY 3320
 QY 2816 TTLCOOQIKTKPLPTNHAEPSPHTNPILOLNOHTQTLTYRPHPTL---ITANTPRQ 2872
 DB 3321 VAVCOQEVQARLIPDYASHSHVEDLKELERVLSGIRSPRPVCSVAGEQGE 3380
 QY 2873 LITPHYTQOARNTVDYATTTQTLHOGCTTYIELQBDNLT---TLTHNLNRPPTTL 2929
 DB 3381 VPDAGVFNLRNVERFSAVVGGLBQGNRFLVSAHPVLAHQTEAADRSVATG 3440
 QY 2930 TLTHNHPOTHLTLNLAK-----TTTTW-----PHYTHNDQPHHTHLDPTYFOH 2980
 DB 3441 TLRQDDSPH-KLITSTAEAMAHGATLMDRALPGHLT-----TLPTPRFN 3487
 QY 2981 HHYLESTQPAQNVSA--GLDPTBRLGATLELATDQALLAGRLSLRSHPWLAHA 3038
 DB 3488 HHYMAVTSRAGVDDAAAGRGMTWEDHPLRGLPLADSGERVAFGRLAGSEHMDLTDHA 3547
 QY 3039 VGGTILSGATFIELAHGTYVGCORVDELTHAPLVYVVDGCVSVQCVVAAADGEGRR 3098
 DB 3548 VSGVTLLPGTAFEFPLHMAAAGCGRLBELSYEARLVLPAAAGVRQVMKSADESGRR 3607
 QY 3099 LVSVYARGSA---CGGAGSGGVMTCHASGVLYVEAAGGVVDGAGVPRGAVADV 3155
 DB 3608 RVAIHSAPEAAVNSAABEGDSAGVTRHGCTIVPRPEPRPADMARAMPAGE-RVER 3666
 QY 3156 DGVRDLRAGACVLGPFVSGRLAVRDGDLLAEVCLPEBAWDAAGFGLHPLLDGVQ 3215
 DB 3667 AEIYERGALGYEYGEAFGVRAVWQRPALLAEVLIPRASYGAGRGVHPLDLAALQ 3266
 QY 3216 PLSVLLPFG-TGEGEGAGFGEVGRVAVWGVSLHBACTGVAVYRISAVRGGRGKREAVSV 3274
 DB 3272 P---WTAGGILEVBEBA-----VLRPAMQGVSLYATGAGALRVRLTKADG---AVSL 3274
 QY 3275 VVGDGAVPVYASVDRLEIRPVMDGOLRA--VSVAGRGSLYAYO-----MAE 3320
 DB 3775 QAADTSGAAVLSGALVMRPLARRKLDVLGTDAGER-SLIRVEMQRLPLPAPRPSWAV 3833
 QY 3321 VGP-----VPVCGQAMAHNEVDVSGGGPV---PGVVLRCPDAGAGG----- 3360

D	b		 LGPDAADRLAGTGGGACD-----QPRSGGTALTYEVRALR--KALLAAGRPREDEAVVLR : :	3862
Oy			--GGGGGGGGGVGEVVGVGLGVVOGMCLERFPASRLVVVTTRCAUVVAAREDPYDVVGAS : :	3418
D	b		VLSGAGATPESVQRQTERTCTLAQDWIDAEELVPRLVLTVIRGAIVAAPSEIIDLACAG : :	3942
Oy			VWGLVRSQAQHNPDRFLVDLDTDTGTLDTLGACAGKVDSGRVAUVAVCSEPQLARVE : :	3478
D	b		VWGIVRSARSSEHPGFALVDTD-----GHPPDDRALT-----PLAIRAV : :	3980
Oy			RLLAAARKLTSSGSQVPAQRSGDTPRA---RRSDVPAQSRSGSVPARRSVDVSGREVLPMLS : : : :	3535
D	b		-LDGAGCL-SLRAGSTATPVYLTLAQTREEERG-PA-----FDF : :	4015
Oy			GGSVLVTGGTGVLCNAVARNHLAGCVARDLVISRRCRDAPCAGSLRAELAAGAERYIV : :	3595
D	b		AGTVLVTCATTGLRIHLARIHDIAEHGVNHHLLLISRGSAAGCADBELMELAGLEAPERFA : :	4075
Oy			ACDVERREVRVRLLEGVPACRPITGVUNHAAGVLDATIASITPERLGTVFPAKYDALLL : :	3655
D	b		ACDAADRREALRVILAEVADPADPRTGVIHAAGVLDGDTDLATLPRI GTVMSPKAADAAML : :	4135
Oy			DELFRGMELSFVLFSSAAGTIGSAGOSNVAAAABALDALAYRRRAAGLPGVISLANKGME : :	3715
D	b		HETRTSPLSAFVAFYSGAAGLIGRPQQANVYAANTFDIALAQHNRAHGLPWSLAWMGWG : :	4195
Oy			EASGMTGLACTDHNRITIRSGIHMPSTEDALFDPALA-LDRPVLLPADLRPA---- : :	3768
D	b		GATGHTGHSSTDRLRMKRSGSIARPYNDOGLALDFRALMAAEBDLVPHMLDALVRE : :	4255
Oy			PPLPYLODLLPATRR-RTRTTTTGGAD--NGOAHARLAGOTHEOOHTTL : :	3818
D	b		RAHQRPADVAPPBLGLIPARAAYROAAPRVAGCARPARAGCGTAERIAGLGEARIRELY : :	4315
Oy			ALVASHLATVUGHTTPDTITPPDRAPRDGPFDSLNAVELRNLSKTGRLPTTLAFPDPNP : :	3878
D	b		RLVARVESGVGVSGPDVAVEGRPCKDGPSLTAVALERNLGAATGIARLFTALVFPBP : :	4375
Oy			PTTTLNHIHOTOPQP-D--NNVAPVLAELDYLESLMDJDTDSASB-----RYTLRLUKSL : :	3932
D	b		GCNAVAEVILAELAPRGDCGDTAAAAFEGLEAANAANGLAEDDLRRDYLRRLTELAAL : :	4435
Oy			MWRMAPQHPTAESADDDEKTSATEAEIFPKRIDNL : :	3969
D	b		TPOGNPESAPA PAPSDULDERLSANDDDLPAFIIEQL : :	4472
RESULT 8				
ID	O93HU5	PRELIMINARY:	PRT,	6146 AA.
AC	O93HU5;			
DT	01-DEC-2001 (TREMBLrel. 19,			Created)
DT	01-DEC-2001 (TREMBLrel. 19,			Last sequence update)
DT	01-JUN-2002 (TREMBLrel. 21,			Last annotation update)
DE	Modular polypeptide synthase.			
DN	OLMAL1.			
OC	Streptomyces avermitilis.			
CC	Bacteria; Firmicutes; Actinobacterii; Actinomycetidae;			
CC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.			
OX	NCBI_TaxID=33903;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21477403; PubMed=11572948;			
RA	Omura S., Ikeda H., Ishikawa J., Hanamono A., Takahashi C.,			
RA	Shinohe M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,			
RA	Kikuchi H., Shiba T., Sakaki Y., Hatcortliff M.;			
RT	"Genome sequence of an industrial microorganism Streptomyces			
RT	avermitilis: Deducing the ability of producing secondary			
RT	metabolites".			
L	Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).			
EMBL	; AB070940; BAB69192.1; -			
OR	InterPro; IPR002106; AntRNA_11gaeeit.			

QY	DB	Query Match	Best Local Similarity	Matches 1849	Conservative 469	Pred. No. 2,86-314	Mismatches 1395	Indels 1181	Gaps
QY	Db	4 MDGSEBRPAAAGVTLGVADELDDGVVVPFGQGFQWGMGRELDDASVFPRESRACCA 63	36.1%	Score 7516.5	DB 2	Length 6146			
QY	Db	1504 LDALAEERTPGVVV-----EGGTAFLFSQGSQIRAGMGRELYSAHFVFAADDEVCA 155T							
QY	Db	64 FAPVVDVSVEQVLRDSDPADGLDDEVVQPTLFVPMISIALMRSQGVPCAVLGHSLGE 123							
Db	1558	FDGLLPGALAEVVE---CGAGLDDTEWAQRLFLALEVALPELVSSWGVRAVLVGHSLGE 16141							
QY	124	IAAAHVSGLSLADAARVVTLMSCAQTTTLACTGALVSAVATPDELLPRIAFTEDNPARL 183							
Db	1615	LAAYVVGWVSLADACRIVAAARGRLMQLPECGAMVVAEADEL-----PDLPGCV 1666							
QY	184	AVAAVNPRESVNVSGAREVAADLVADLTAAQVTRMTIPVDVPAISPLTYALEERVSSLL 243							
Db	1667	SVAAVNPRESLVLSGDEEPYALAAQAPAEQGRTRKQTLVSHAFSHARMEPLAEFTTLA 1726							
QY	244	PITRPSRIPRPHSSSVTSGRLDTRLEDAAYVRRNSSITVRFEPARLLLQOGPKTFVENS 303							
Db	1727	AVERPRPRIPVSVNTGEMAGDEFTTFAVYVHRVREAVRFADGVCAHAAGVRRYLELGP 1786							
QY	304	HPVLT----- 308							
Db	1787	DGLTTLAHDPAHDAGRDVAAPALRRDDQEPETRTVTVAVGRVLYATGLTVDMAPFAPYGA 18466							
QY	309	-----MGIOELAP----- 316							
Db	1847	RPELPTYAQHRRYLEPVAFSSAPRANIPRYRTANASVPDGRPSLSGTMAVVVPASG 19060							
QY	317	-----DLGPT----- 321							
Db	1907	SYEDIYAGVTAAALRDHGAPRVVVEVDAPGERADVADTLRALALGEDAGGVVSLTGASL 19666							
QY	322	-TGTADTVI-----MG----- 331							
Db	1967	SSGVAAATRAVLYQAWTDLGGGARLMCLTRGAVSVSPSDPLAAPAQALMGKRVALLHRE 20266							
QY	332	----- 331							
Db	2027	RMGGLVDLPAVPPDDRAWTRLCAVLAGSGEDQAVRAAGLFARLGRAGTRADAAPAGPS 20866							
QY	332	-----TLRG----- 336							
Db	2087	PWRDGVTLVTGGTGAICAHLMQWLAAGAANHLITTSRRGEDADGAABELTARLRETGV 21466							
QY	337	-----QGTLD----- 341							

```

Db 2147 TAAACDVTRDALAEILLSGLPADRPILGVNHAAGVLDGVLDSLTPDRFAAVAGPKVTGA 2206
Qy 342 ---HFLT----- 345
Db 2207 WHHELTFRDLDLSAFVLSSFAQTVGLAGQANVAANAAYLDALAVHRAQLGLPATAVAMG 2266
Qy 346 -----SLAQLRGHGETSATP----- 360
Db 2267 SMSGAMMAADTEARROLARTGLVPLAPDALALALGOVMDGGETRANVYADVDMDRFAAGF 2326
Qy 361 -----VLSARLTALSPTQOQSLDLIVRAHTMAVLN 391
Db 2327 GPGRPSPLITGVPEVRRIEERETGRAGAPAEITRGLTGLTALQALADLVSTVAAYLIG 2386
Qy 392 DDGHERTASDAGPSASFAHIGFDSVMGVELENRLSKATGLPLPTLLPDHTTPAAVAVARL 451
Db 2387 LPTDRVED---TTFRALGFDSLIGVEFRNRLLAAATGRLLPSSLFPDHTPGRLEVENL 2442
Qy 452 RTAALGHLEBDTAAPVPSGSGHTAA---DDPIAITGMACRFPGVSPKDLMELAS 508
Db 2443 AAGLDG-----TGF---NTAGEERVAPARTHDDPVIVSAACRFGCVTTBEIMQLVLD 2494
Qy 509 GGDALGPPPTDRGMPTQGRHAODPTQPGTFYPOGGGFLHDAHFADAGEFGISPREALAMD 568
Db 2495 GGDALGPPPADRGMDLRLYDPDGNPGTSYVRECGFLTVDVADFAAFGISPREALAMD 2554
Qy 569 POQULLETSWEAREBAGIDPLSVRGSTGVFAGALSDFYGRMDTASBEAAVEGHIL 628
Db 2555 POQULLETSWEALERAGIVASLASRTGVFVSGNGDY-----ATLLGESEBEGLVL 2608
Qy 629 TGTGTVLSGRIAVSFGLGEPATITVDTCASALVTLHLACOSLSRGECTLLAGVSVMS 688
Db 2609 TGTASVLSGRIATTLGLEBALTITVDTCSSSLVALHLAYQALNAGBEDLALAAQVYMS 2668
Qy 689 TLGMFIEFSRQGLSVDRCKAYSAADGTGMEGCVMLVERLSDAVRLGHRVLAVERG 748
Db 2669 SPEIFVEFSRQGLSADRCRCAFPGDADGTGMEGCVMLVERLSDAVRLGHRVLAVERG 2728
Qy 749 SANVODGASNLVTPNGRBOERVLROLANAGLSVANVDVVEGHTGTTGDDPIEAQALL 808
Db 2729 TANVODGASNLSPANGPAQOQVIRQALADAGCVSPDVAEALGTGRLGDLPIEAQALL 2788
Qy 809 ATTYCO-RAGDRPLMGLSKSNIGHTMAAGVGVKJMWALREGLPTLLHVDSPQVD 867
Db 2789 ATTYGDRPADRPMLGSKSNIGHTMAAGVGVKJMWALREGLPTLLHVDSPQVD 2848
Qy 868 WSAGAVRLTEAVFWPGDAGRLRAGVSSFGIGTNAHVILLEBAPAGGCVAGGVLEG 927
Db 2849 WTEGAVELTETREMPATGA-RPRRAGVSAFGVGTNAHVILEQAPAP----- 2895
Qy 928 APGLAISVABSVAPVAVASVAVSVPPVPPVPSVASEBGLAQALAOAYAVR 987
Db 2896 -----OPARPAVLPEPLNSP-----SVPMVLTAKESEBALAAQADRLS---ARL 2935
Qy 988 PDVSLADVAGIACGRAVLLEBRAVVLADRELEVQGLALAAAGEPDRVTGHAPEGDRG 1047
Db 2936 PEAPARAEVGRSLVTRATWPERAVULIGSGPHPAAGLALGAGCDPRVVRGTA--DTRG 2993
Qy 1048 GVEVVFPOGGQOMAGMVRLLASSPVPARMQACEALAPVWDSVVDILRRDAGDAWE 1107
Db 2994 RVEVVFPOGGQOMAGMARLWESSPVPFQOMDRCEVLAARLTWDSLVEVIHETODAPOLD 3053
Qy 1108 RADVVOPLFESVMTSLALMRSYGIERPAVLGHSGOGEIAAAHVCGALSLKQAAKTVALRS 1167
Db 3054 RVDVVOPLFESVMTSLALMRSYGIERPAVLGHSGOGEIAAAHVCGALSLKQAAKTVALRS 3113
Qy 1168 RALA-AVRGRGMAVPL-PAQEVOLIGERWAGRLWAAVNGRSTAVSGDAVANDVL 1225
Db 3114 RLIEBELSGHGMVSVALSPADTART--ARMNGICVAANNRSTVSVSEPPALABEL 3171
Qy 1226 AYCAGTVRARIRIPVDYASHCHPHVQPLREBELLELIGDISIPQSGVPEFSTVEGTMDTJT 1285

```

```

Db 3172 ABCEAEVGRARIRIPVDYASHSAOVERIERQLTIELAAGVAPRGDIPFHSITTTGGPLDTTG 3231
Qy 1286 LDAAYVYRNHOPVRFSDAVOALADGHRVFEVSPHPTLVPALEDTTEPTAEVDVAIGS 1345
Db 3232 LDADYTRNLRPRVLRPPTVEELAQGHVFLMSHPVLVPAVQADAVDATGTTAAVGS 3291
Qy 1346 LRRGNDTRRRLTALAHTHTTGTIGTPTTWHNHYTHHTHHPHT-----HLDLPTY 1396
Db 3292 LRRDGGPERFTLSLAEAFVR--GAPVDM-----TAVLGTSARRVELPTY 3335
Qy 1397 PFGQHWTLSQPGAGSGSAGAGSAGAGTAGTAEVESRFPDADARODLEAVAT 1456
Db 3336 PFGQRFMPERPT-----AAATGG-AEBDAPLWRAVDRGDLAAVAA 3376
Qy 1457 TLAVPESAG---LDVVPALSAHRRHOQARINTWYQETWKEPLTPTTHQPHQWLI- 1512
Db 3377 ELALPQESLHDLDVLPALSGRRKRKDALVDMSRYKRSWTPVAPVLAALTGWMLV 3436
Qy 1513 -AIPETQTH-----PHITNLTMLHNGITPPIPLTNHTHTNPQHLHTHTHQQA 1564
Db 3437 TADPDGEAMARATLGAATQUTSPD---GPFP----- 3468
Qy 1565 QNHITGAILGLSLALADETPHHPHTPTGTLNLTLTQTHQTHPTPLMTATNAT 1624
Db 3469 ---AGWTVGSILGLADRPBHPVVPVAVADVALLSALREAGIDAEPLMGTRGAVG 3524
Qy 1625 THPNDPLETPHOAQTWGLARTLLEHPHTHAGIIDPTPTPTPHLQTLQTLPHHQTO 1684
Db 3525 TGRSDOVRAGQOLMGLGVAGLENPAMGVLVDPAELDDBAAGILBALAAEBEOE 3584
Qy 1685 LAIRTGTHTRRL--TPTTLTPHOPTPTPHGTTILTGSGALATHLTHLTHHOPTH 1742
Db 3585 YALRPAGLYVRRLVRAPLDGTARBSWRPRPDGVVVTGATGALGARVARWL-AHAGAGH 3643
Qy 1743 LLTSRGTGPHTPHQAHTTLOQKGIHLLTTTCDTSNPDLOQLNTIPRQHPLTYIHT 1802
Db 3644 LLTSRSGMAADAGDLDVLAELRAGVAVAACDVADRAQALADLVDPDEFVAAVHT 3703
Qy 1803 AGILDDATLTNLPTOLNVLRAKASAHLLHOLQHTPTPLTAFLVSSAATGAPGOAN 1862
Db 3704 AGVNRKTTLRRTLLAEIADVMSAKAGARULDELACHDADAVLPSGAILNGSAGQA 3763
Qy 1863 YAAANAYLDALAHRRTHLPLATSIWGTQNGLADSDKARAYLDRGRFMSPELATA 1922
Db 3764 YAAANAYLDALADRRRRLGATSVAMSGAGGMVDDLDAR-LARGVRSMDBRAIG 3822
Qy 1923 AVTQALDTERPVYVADIDMSKIEHTSOTSIDVSAARER-----EPVQRPPTPA-- 1973
Db 3823 ALQOAL-DHDETLTVTDMWERFAGT-----FTIARPRPLIDAIPEARAAATTPBAG 3875
Qy 1974 --ELHKTLAHQTSADQALLELVRDHVAVLBHADPKAIAPOSPRALGFDLSITAVEFR 2031
Db 3876 ESELAARLASTIGERRELELDVTRPAALALGHRGAELTTPSKPFODLFDLSITAVDLR 3935
Qy 2032 NLLIKATGLRPVSLVFDHPTPAKLAVHLONOLRGTA--ESAPSAAVTAASVTEPIA 2089
Db 3936 NRLTAANTGVALPATVFDHPTPLAAVELTRLPFGGALPTSAGSVAPTAABD--DEVV 3993
Qy 2090 IYMACRFPBGVTSADPFMDLISSEBODJIGCFPTDRGMDLTYLDPDPHPCGYTRNGG 2149
Db 3994 IYMACRFPBGVTSADPFMDLISSEBODJIGCFPTDRGMDLTYLDPDPHPCGYTRNGG 2209
Qy 2150 FLVDAGHGFDEFFGISPREALAMPPOORLLETRAMETIEBAGINPTLHGTPGVFGTN 2209
Db 4045 FLVDVAFEDGFGISPREALAMPPOORLLETRAMETIEBAGINPTLHGTPGVFGTN 4104
Qy 2210 GQDYALRVNNAAGOSTGFALTGTAGSYISGRISYTFEGEBPAYSVDTACSSSLVALHLAC 2269
Db 4105 SHDYGTLTSLBEGOD-YALTGAVGSVLSGRILVYGLBEPALTVDPAACSSSLVALHLAA 4163
Qy 2270 QALRAGECSMALAGTGVWSSPAFVEFSQRLGADGHCAPSAADGTGMBGCVGML 2329
Db 4164 QALRAGECSVALAGVAVMTAPDAFAFARQGLADPGRCKAPADGADGTGMBGCVGVLV 4223

```

QY	2330	IERISDARNHRVLAUVRGSAVUNDDGNSGTLTANGSOORVURQALANGLSAGDVDA	2308
Db	4224	LTRISEARRKHGDVLAUVRGSAVUNSDGNSGTLTANGPSOORVURQALASGSLAADVDM	4283
QY	2390	VEANGTGTLLDPLEAQALLATLYCQDRAEGEGLTMSYKSNVGHITQAAAGVAVIYUWMA	2449
Db	4284	VEANGTGTSLDDPLEAQALLATLYCGDRADRRLMIGSYKSNIGHITQAAAGVAVIKSVLA	4343
QY	2450	LRLHGLLPTTLHAVDESPHYNDNSAGAVOLLTEVPMRGEGRLRRAVGSFGVSGTNNAVY	2509
Db	4344	LRLNGLLPTTLHAVDESPREVDWSAGAVELLTEGREMPEDGR-RRAGVSAFGISGTNAVYI	4402
QY	2510	LEAPRADVPCGPRRGEGDAGSDDEBAAGSPCWMMVLSAKSOPRLRYAOAL--HANYL	2567
Db	4403	LEQAPSVERTAPRADG-----LIVPMVVSASSEELRYAOARLDHVAAT	4449
QY	2568	DHPGLDLADVGYTLNABAVRDNHRTLAAADRFTGLQALQALAAGERPRAVHSSAPGT	2627
Db	4450	D---LTVADVUSLSTAGLEHRAVULTEGREDEPLVQALALAEQ-----ASAAMV	4498
QY	2628	GTEGAAGTATIGSGOQORPGMAHGLNHTPRVAALNDSTHLDHLDHRLYRLTQN	2687
Db	4499	RGVAGEGTAALFTGOQARLQSMRELERPRVFAAADGVCAODLLEERUVEVUTD	4558
QY	2688	DNDNEADALLQOTTEUQARLFAFOVALNRLTDOXNHTPRYUAGHSIGETLAAISL	2747
Db	4559	-----ARALDRITFTQAGLEFALEVALFELVS-SMGVADVLLGHSIGELAAVAVGW	4610
QY	2748	TLTDAATLLITOPATLMQMPRTGTMTLTTPHHITNLHANENDL-----ALAIATP	2800
Db	4611	SLAASCRVVAARGMLQALREGBANVA-----VEAADEBLRDLREGSVVAANVR	4660
QY	2801	TSLVSGSRHTYONHTLCSOOGKIKTKLPTMNAHSPHTPRLYQNLQHTLQTLVYHPR	2866
Db	4661	RSVLUSGEBEVUTALARTPAEOGRKTKDLAVSHARMSRMYLAERETLAAVEFRPR	4720
QY	2861	TPV---ITANTPRDOLLTPHWTQOARNTVDYATTTQTLNOGVUTYLEGRDNTLTTL	2917
Db	4721	IPVVSNTVGMAGAEFTTPRAVWVRHVRBAVPAODAVLVARGVDFLEGRGALGMA	4780
QY	2918	HNHLRN--PRPTTLTLTPHHNPOHNLTLNAKTTTNNHNYTPHNDOHTPHNLDERT	2975
Db	4781	ERTLDHGTADAVCVVULHRREREDATLLNALATASVTAADVMTTPRACAT--TELRTP	4837
QY	2976	YRFQHNHW-LESTOPGAGNVSAAGLDTEHRLGATLELATDGCALLAGRLSLRSHWL	3034
Db	4838	YAFQRRRRWPRRAASAGADLAAAGLAAAGNRLTLWMLPRSEGVDSULTGRLSLATNHWL	4897
QY	3035	ADHAVGTVLLSGATFLELALHAGTVVCCDRVDELTLPARLVVPVDSVSVQVGVAAADG	3094
Db	4898	GDHVTALGTVLPRGTAFVULVGMVGBRGVCCGIRELTLPATRLAEDAVLRVLYUGABDD	4957
QY	3095	EGRRLVSVYAGKSGSCGGGASGGVWTCHASGVLEAAG-----GVVVGLAGVWRP	3147
Db	4958	TGCRPVAYAPDTE-----DEGDMTHAEGELARADASTOPRAPRPAID---AMPV	5008
QY	3148	RGAAVADVGVDRILAGACVLYVSGRLAVWRDGDLLAEVCSUREAMGDAGFGHPR	3207
Db	5009	PGAERPLPMDGFEGLADAGPAVGRFLRGLRAAMRGDEVPAVSLPRADESG--AGFGVNR	5066
QY	3208	ALLGCVQPRVLVLLPGCTGFEGSAGFGSGCVVPAVMGVVSLHRYAGVTCVPRVAVAGVGG	3267
Db	5067	ALLDAAALALG---PRAQASDEPRG---SALRPFSGVGRVNAHAGADLLKRLVLYAADG-	5118
QY	3268	GRAVSVVVGDEAGVPRVAVDRLBELRPMVQOLRAVVSASGRGSLYUAVQWAEVQPRVVC	3327
Db	5119	---TUTLDAADAAGRPVVSIGSLVLRPLSPRLRAGTQAPD--DALFATRMV---PLDVT	5170
QY	3328	GO-----ANAMHEDVGE-----SGGVRPRGVVULRCRPDAGAGGG	3362
Db	5171	GDUTGATADVECVLLGSPSAGAMKSHDLAAUAAALADGKETPRVLAERPLT-----	5224

OY		3363	GGGGGGGVEVGVGLVCGVMGLKBPASRLVVYVR--GAUVAAGDEDPDVUGA----	3417
Dd		5225	TGDPAEARRSAEWTLIDLQOMWLADDBLTSHLVTGRHAAPVATPYAAPVDATADDAADP	5284
OY		3418	-----SWMGLVSRAQAENHPFVLIDLDPTGTDLDTGCAGAGCVD	3458
Dd		5285	VDACHENAPVDTPAAAALAAQLSTLAGIVRSACQENPARITLVDFDT-----APD	5331
OY		3459	GGRVAAVVAAGEPOLAVRGERLLAARLK-----RLESSG-----	3492
Dd		5332	PAHLTAARVLGEPEVAVRQGTLVARLRTRPETGRALTVPRCGPBMRLDSTGRGSLDHIAL	5391
OY		3493	-----DVARQRSG-----	3500
Dd		5392	VPCPDAAAPLAEGTVRIAVAAGVNFVDYLIALDMYPGRADLGTECAGVVLSETGPGVTGL	5451
OY		3501	-----	3500
Dd		5452	APGRVWGMAGAAGCPFAVADVRLAARIPOKSEFTAAIPVAFLTYVGLVDLAGIAG	5511
OY		3501	-----DTRARSVDVAOR-----	3513
Dd		5512	ESVLVHAAGVGMAAQOLARHLGAEVFGTASEPKMTLLDYGDRAHICSSRTLDPADL	5571
OY		3514	----SGGV-----PAR-----RS	3522
Dd		5572	VRETTTGAGVUDVVLNSLAGEFTDASLRAMPBGCRFVEMGKTDPDRPERVAADHPGVYRS	5631
OY		3523	VD-----VSGREVL-----	3531
Dd		5632	FDLBAGRPDIRIAETLANLVELFBESGALTRYLPVTTWDIRATDAFRVLSQATLTGKAALTV	5691
OY		3532	-----PWLSGGSVLVTGSGVLGAANVARHLAGVCVBDLLVSRKRGDAPGAEGRAYELA	3586
Dd		5692	PAGSVPREAGETVLTIGGTGTLGTLARHLVTEHGRLNHTLAGRGDTPDEVNQLRADLA	5751
OY		3587	ALCAEVRIIVADCVEREREAURLLEGVAPGECRLTGVUNAAGVLDATTASLTPEPLGTVFA	3646
Dd		5752	EMGAEEVVEACDADDEDVAVRLLLEPLTAHHRLAGVNAAAGVTDGCVVASLDRRLSYVLH	5811
OY		3647	AKVDAALLBELTTCGMBLSAFVLFFSSAAGILGSAGCNVANAANALDALVYRRRAAGLPG	3706
Dd		5812	PKVAGANNLHRILTAMHLPRMFVLFPSSASATLGAGCGNVYAANAFDALLBHNRARCLPA	5871
OY		3707	VSLAMGLEMEBASGMTGHLAGTDHRRILIRSGLHPMSTPDALDPAALADLRPVLLPADLR	3766
Dd		5872	SSLAMGLMDQASSGTGTLAREDRQMRBAQVAAPLDSANGALDPTALLTGIPALTPEARLD	5931
OY		3767	PA-----PRLEPLLQDLLPATRRRTTCT--TTGADNGAQLNARLAGOTHEOHTT	3816
Dd		5932	LAGRABARAQPVAUVAVRELVRVARVAATGTAPTGTGPDARLLTERLLASPABRHRL	5991
OY		3817	LILAUREHIATVUGHTTPDITPPRARARDGFOSLTAVELRNLRSRTTGLRLPTTLAFDH	3876
Dd		5992	VLDLVREHTAGVLGHSADDIIDPOQAKTKGFSPLTAVELRNLRKRTASTLVPATTVFDH	6051
OY		3877	PNPTTLTHNHITLOLOPQDNAAVAPVLAELDKTSALSALDKTDSASEBYTRLKLKSLRW	3936
Dd		6052	PTPALATLPHLELLELAPREHDPVLOVGLERLERLEGVALCSPVDGANDVAVARLRIRIKA	6111
OY		3937	NAPQHPTAESADDEKFTSATFEAIFKFINDDLC	3970
Dd		6112	EAGSAP--DGAGEDSSLOTASAABVLAFTDSERG	6143

RESULT 9

O93HU1 PRELIMINARY; PRT; 3613 AA.

AC O93HU1,

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Modular polypeptide syntheae.


```

Db      1550 SATETLPLVLRRTD-----DEIVYMGACRPGGVRGPEBDMRLVDSREMDPFRADRG 1605
Qy      2127 WDDTLVDPPDHGTCYTRNNGFLYDAGHFDAAFFGISPREALANDPOORLLLETAMET 2186
Db      1606 MHLGAMMAFIEESG---AAGAFIABAGFDAAFFGISPREALANDPOORLLLETISMEA 1662
Qy      2187 IEHAGINPHTLHGTPGVFTGTNGQDYALRVHNAAGSTGDFALYGTAGSVISGRISTFG 2246
Db      1663 LERAGFPVGLRGSRGTGVFGTFQBYETVLVMSSEAGGGVALLTGASGVMSGRVAVALG 1722
Qy      2247 FEGPNAVVDTRACSSSLVALALACALRAGCSMALAGCVTVMSGPCAFVFPESRORGLAAD 2306
Db      1723 LERPAVAVDTRACSSSLVALALACALRAGCSMALAGCVTVMSGPCAFVFPESRORGLAAD 1782
Qy      2307 GHKAFSAADGTGWMGVMGLVERLSDAHNRHRYLAIVRGSANVDDASNGLTAPNG 2366
Db      1783 GRCKAFPAAGDGTGMBGVMGLVORLSDAVDRPVLAVRGSANVDDASNGLTAPNG 1842
Qy      2367 PSQORVIRQALANAGSAGDVDAVEAHGTGTLGDPTEAQAALATYQODRAGGRLWLS 2426
Db      1843 PSQORVIRQALANAGSAGDVDAVEAHGTGTLGDPTEAQAALATYQODRPAORPLWLS 1902
Qy      2427 VKSNVGTQAAAGVAGYIKVMALRHGLPRTLHDESPHVDMSAGAVOLLITETVPWG 2486
Db      1903 VKSNVGTQAAAGVAGYIKVMALRHGLPRTLHDESPHVDMSAGAVOLLITETVPWG 1962
Qy      2487 GEBRLRAGVSFSGSTNNAHVLEBARADVDGCRPAGSGDASDEDAAGSPGVWML 2546
Db      1963 TDBP-RKAGVSFSGSTNNAHVLEBARADVDGCRPAGSGDASDEDAAGSPGVWML 2008
Qy      2547 VSASQORVIRQALANAGSAGDVDAVEAHGTGTLGDPTEAQAALATYQODRAGGRLWLS 2604
Db      2009 VSASQORVIRQALANAGSAGDVDAVEAHGTGTLGDPTEAQAALATYQODRPAORPLWLS 2065
Qy      2605 ALQALLAAGEPHAVIHSSAFSGTGTGPAAGKTAFISGGGTORPMANGLYHTHPVFAA 2664
Db      2066 QLAALLEGAGAGVGRVA-----GE--GTAFLFTGGCAQRIGMRELYERFPVFAA 2117
Qy      2665 LNDICTLDPHLDHPLRLTQNDNDNEAALLQORVAPORLPAFOVALHRLDTCYN 2724
Db      2118 FDOVCAQDLLEPRVCEVFTD-----APALDRVFTQAGFLAEVALLFELV-GSWG 2169
Qy      2725 ITRHYVAGSHSGLTAAHLAGLITLTDATLITOTRATIMOTMPR-GTMITLHTPHNITH 2783
Db      2170 VRADVLLGSHSGLTAAHLAGLITLTDATLITOTRATIMOTMPR-GTMITLHTPHNITH 2225
Qy      2784 HLTANENDLAAIINTFTSLVIGTPHTVCHITTLCOQOSIKTKTLEPTNAPHSPTNPI 2843
Db      2226 ELFPVLPGSVAAVNGPRSLVLSGDESPYTLAQAFAEGGRRTQOLAVSHAFSARMEPM 2285
Qy      2844 LNLQNHOTLTLNHPRTL---ITANTPRDQLTHTVHTQOARNTVDVATTTQOLHNG 2900
Db      2286 LAFERAEFLAAVERPRRIPVASVNTGGMADAEFTTPAVVWNRHREARFADGVAIVLARG 2345
Qy      2901 VTTVIEIGRPNTLTLTNHMLPN--PTTTLTLNPHNHPOTLTL-NLAKTTTWNPHN 2957
Db      2346 VDRFLELGRGALTAAMEETLDTHTGADAVCVPLHP-ERERADSLITALGRHIASGAPVD 2404
Qy      2958 YTHNDNORHTHTLDTLTPFROHNHNLSTORPAGNVSAAGLDTEHPRLGATLBIATD 3017
Db      2405 WSGLPAGTGART-VGLPTVAFONHRYVLSRTTSGSDPASIGLTATGHPRLGAGVALPDS 2463
Qy      3018 GGALLGRSLRSHNPLADAVGTVLISGATFELALNAGTVVSGORVDELTHAPLV 3077
Db      2464 DGLFTGRSLATOPWITTDHMTGTLTPTAFELALNAGTVVSGORVDELTHAPLV 2523
Qy      3078 PVDGVSVOVVAADDEGRRLVSVVYARSGSACGGGASGAGVTVCHASGVVVAAGGV 3137
Db      2524 GEGGGRVVOVLVGLDGTGRATLHSH-----PDGGDDQPMILRHASGVLAENTGAPD 2578
Qy      3138 VDLGAGVPRGAVAVDVGDVDRBLAGAGCVLGFVPEGLRAVWRDGGDLAEVCLPEAV 3197

```

```

Db      2579 AEPMT-VMPDGGSTAVAVEDFPDMAEAGFTYGPVFGILRALMTKDGLVAVRLBEBA- 2636
Qy      3198 GDA-AGFGLHPALLDGVQPLSVLPGCTGFGAGFGEGBVRVPAVMGVSILHAGVTG 3256
Db      2637 GDDSGFVHPALLDAGPLALGVGTGTPDRANVCG---MPPAMTGRILATNATVA 2692
Qy      3257 RRVNSAVRGGGREAVSVVGBEAGVPVASVDBLELRPYVMQGLRAVSVSAGRGSLYAV 3316
Db      2693 RYHILAEVGRG---EVSVLVNBESGLPIATVGSILTRDPLBQPTA---SAPQDMLFDL 2745
Qy      3317 QMAEVPVP-----VCGQAMAM-----HEVGESG-----GGPV 3345
Db      2746 KMT---FVPLSGSERVSGE-WAMIGFDPLEIRPLVAGLTGTPPYLDPOSILDTVESGKPA 2801
Qy      3346 PGVVVLRCPDAGAGGGGGGGGVEVVGVLGVVQGMVGLERFAGSRILVVVTRGAVVA 3405
Db      2802 PSVIVAMSC-----FGDDGGGIYAATIDAVHRYLVEVQHLIADRLTGSKILLITRQAVPA 2856
Qy      3406 GPEDEGVDVVGASVWGLVRSQAENHDPFVLLDLDITGTDLDITGAGAGVNDGGRVAAV 3465
Db      2857 AGTDRTEVDVAASAVWGLVTRTASBHPDRIVLIDDDPAS-----YRALPKA 2903
Qy      3466 VACGEQLAVERGRLAALAKLESVDYPAQRSGDTRARRSDVPQRSGGVARRSVDV 3525
Db      2904 LGTGEQQLALRAGSPCVPRILARHTGPAD-----GAPG----- 2935
Qy      3526 SGREVLPMWLSGGSVLTVGGTGVGAAVARHLAGVCGVRLDLVSRRGDPAFGAGLRABL 3585
Db      2936 -----FGPDGTVLTGGTIGALGAVVARNHLYTHAGVRHVLVLAGRGERPAGAAALLAEL 2988
Qy      3586 AALGAEVRIACVCGERRERVRVLEEGVPAGCPLTGVVHAAGVLDATIASLPERLGTVE 3645
Db      2989 TELGADTRILACVSDVRDALAVLIRDI PADRPITAVHAGVLAGDTSVSLTDRDVTYL 3048
Qy      3646 AAKVDALLILDETRKMEISAPVLFSSAAGILCSAGCGYAAANALDALATYRRRAAGLP 3705
Db      3049 RAKVDAAHMLHBTDPARLREFVLFPSSAAGLGSQCGYVAAANSFLDILAAHRRREGLP 3108
Qy      3706 GVLAMGLMEBASGMGHLAGTDHRIIRSGLHPMSPTPALFLFDAALLDRPVLL----- 3761
Db      3109 GTALANGWMEQFGMGKDLGRABRARMAGGVTPTFAESGMEAFDRLAGGALLVPMR 3168
Qy      3762 ----PADLRPARPLRPLDILPATRRRTTITTTGGADNGAQLNARLAGQTHQOHTTL 3817
Db      3169 LMTTARAASSEQVREPLRGLVAPORRAARSQVRA-----TSHLHQLLAAQDAERQALL 3224
Qy      3818 LALVSHIATVYGHTPPTTPEDRAFDLGFDSLAVELRNLSRTTGLRPTTLAPDHP 3877
Db      3225 ABLIRGEVAVQVGHSGAEALIEBQHFVEIGFDSLTSVELRNLRINERTGLRLASTVFFDHP 3284
Qy      3878 NPPTTLTHHHTOL-QPQPDNAVAPVLA 3904
Db      3285 TPDLALAEISBQLGTPRAAPAAAPAGQ 3312

RESULT 10
ID Q9ALM2 PRELIMINARY; PRT; 5588 AA.
AC Q9ALM2;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Polyketide synthase extender modules 8-10.
GN SPNE.
OS Saccharopolyspora erythraea.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;
OC Saccharopolyspora.
OX NCBI_TaxId=60894;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21257765; PubMed=11358695;
RA Waldron C., Matsushima P., Rosteck P.R. Jr., Broughton M.C.,

```

RA Turner J., Madduri K., Crawford K.P., Merlo D.J., Baltz R.H.;
 RT "Cloning and analysis of the spinosad biosynthetic cluster of
 Saccharopolyspora spinosa.";
 RL Chem. Biol. 8:487-499(2001).
 DR EMBL: AY007564; AAC23262.1; -.
 DR HSSP: P25715; IMLA.
 DR InterPro: IPR001227; Ac transferase.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Peptidase_attach.
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00698; Acyl_transf. 3.
 DR Pfam: PF00109; ketoacyl-synt. 3.
 DR Pfam: PF02801; ketoacyl-synt_C_3.
 DR Pfam: PF00550; pp-binding_3.
 DR Pfam: PF00975; Thioesterase; 1.
 DR PROSITE: PS00075; ACP_DOMAIN; 3.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 3.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; 2.
 KM Phosphopantetheine; Transferase.
 SQ SEQUENCE 5588 AA; 585629 MW; C47F38F44BAC8A11 CRC64;

Query Match 35.0%; Score 7281.5; DB 2; Length 5588;
 Best Local Similarity 36.1%; Pred. No. 3,1e-304;
 Matches 1821; Conservative 490; Mismatches 1349; Indels 1389; Gaps 86;

QY 12 PAAEVLGVADDEADGCVFPPGPGQWPGMGRRLDASDFRSVACFAAFAPYDWS 71
 DB 546 PAAVNVGGAAD-CKGKVAFFPPGGSHMGARELSSEPFVRKLAECACAAATPYDWS 604
 QY 72 VEQVLRDSPAPGDRVDVVOPTLFAVMSIALMRSQGVPCAVLGHSIGEIAAAVSG 131
 DB 605 LGLRLRDPDPAALDRDVTQLAFAMVMSLAELMRSCGVPRAVVGHSCEIAAAVAG 664
 QY 132 GLSLDAARVVTMSQAQTLTLAGALVSVAATPDELLPRIAPWTEDEPARLAAVANGP 191
 DB 665 ALSLTDVARIILAACDAVSALTGKGMIALALPESAVVKRIAGLPE-----LTVAAVNGP 719
 QY 192 RSTVSGARBAVVDVADLTAAQVTRMIPVDVPAHSPLMVAIERVVSGLPITRPSR 251
 DB 720 GSTVSGSEPSALRLQTELTAENVQTRVGIDVASHSPQIAQVGRLLDRGEGSSPAE 779
 QY 252 IPRFSTVGTGRDLTRELDAAWYVNMSTVFEFPAARLLLOQGPKEVMSPHPLTGL 311
 DB 780 IAFSTYTGERTDGRDLADQVYQNLQPRFQGTVAARMADQGRFVEVSPHPLTAGI 839
 QY 312 QELAPDLGDTTGTADT--VIMGTLRQQTLDHFLTSLA--QLRG----- 352
 DB 840 QE-----TLEADAGGVVVGSLRGGGSRRLTSLACQVGLPVMNQVPLNTGAR 892
 QY 353 ----- 352
 DB 893 RVPLPTVPFORQRYWLESSEYDAGLGSVGLLSAENHPLGAAYTLADAGGLLTGKLSVK 952
 QY 353 ----- 352
 DB 953 TQPLADHVVGALLPETAVENTLIRAAQVGCGLLELSLTTPLVLPATGAVQVQIAV 1012
 QY 353 ----- 352
 DB 1013 GGPDEAGRSVRVHSCRDPAVPODSMTCHATGTLTSSDHQDAGGPGGIWPPNDAAVPL 1072
 QY 353 -----HGE----- 355
 DB 1073 DSFYAARAERGFDGPAFOGLQAAMKRGDEIFAENVLPFAHREDAGRGHIALDALAQ 1132
 QY 356 -----TSATTVLS--- 363
 DB 1133 ALGAAEEDPDDEGLPFAAQGVSLKATGALSRLVHLVAGANAVSVFTTDTTGAQVLSIDS 1192
 QY 364 ----- 363

DB 1193 LVLRQISDKQLAARAMEHESLFRVDWKRLSPGAAPVSWAVTGNDELARACGSLGTEL 1252
 QY 364 ----- 363
 DB 1253 HPDLTGLADPPVPPVPPGASQODLDVASSEAAATQRMULDLIQDMLAARFAGSLVYV 1312
 QY 364 ----- 363
 DB 1313 TCGAASGPAGVSDLVHAASWGLNLSAQSENPDRFVLVDVQGTASWRALAAVSGEP 1372
 QY 364 ----- 363
 DB 1373 QALRAGEVRVPRILARCVAEDSRIPVPGADGTVLISGTTGLGLVAREHVAERGVRL 1432
 QY 364 ----- 363
 DB 1433 VLAGRKGMSAPVTDVLDELVGGAAVEVASCDVGRQAQDLRLTTISAEFPLRGVYHAA 1492
 QY 364 ----- 363
 DB 1493 GALADGVESLTPENHAVKVGPKAAGMHLHETLIDLDSFVLFSSFGVAGAAGQNV 1552
 QY 364 ----- 363
 DB 1553 AAANAFLDGLAQRRTAGLPAVSLAWGLEQPSGMTGALDAAGRSRIARTNPMASAPDL 1612
 QY 364 ----- 363
 DB 1613 RLEEMAFRVGSESLVPVHVVDLNALPADAADGVPLLRDLVPAVPRRSVNESADVNGL 1672
 QY 364 -ARLTALSPTOOQSLLDLVRAHTMAVINDDGNERTASDAGPSAPHLGFDSEVMEYLR 422
 DB 1673 VGRRLRLPDQGTQLGLVREHVSATLGHSG---AVEVGADARFRDLGFDLSGVFR 1728
 QY 423 NRLSKATGLRLPTTLFDHTTPAAVAAKRLTALGHLDEBTAVPPSPSGHGTAAADDP 482
 DB 1729 NRLGVLGVRLLPATVAFEDYPTPRALVRLDLKILIGVEAFT-PAPAFAVA---AVTADD 1783
 QY 483 IATIGMACRPPGVSRPKDLMELASGDAIGFPTRDGMPTBQRAADPTQGTFFPG 542
 DB 1784 VIVVMGCRYPGVSSPEELMRVAGGLDAVAEPDDRGRDQGLDPDDRGLGTSYVCE 1843
 QY 543 GGFLLDAHFDAGFFGISPREALAMPQORLLLETSMEAFERAGIDPLSVRSGRTGVFAG 602
 DB 1844 GGLRLDAAEEDAFFGISPREALAMPQORLLLEVMETVERAGIDPLSLRGSRTGVFAG 1903
 QY 603 ALSFDYGPMDTASSGADAVEGHITLTGTSVLSGRVASFELBEPATVVDTCGASLV 662
 DB 1904 LMHHYDGAERTRAPBG---FEGYLGNGSGAGVFSGRVASFEGEPAPVAVTDACSSSLV 1960
 QY 663 TLHLACOSLRSGECTLALAGVSVMTLGMFIEFSRORGLSVGRCKAVSAADGTGWE 722
 DB 1961 ALHLAQALRSGBCDLALAGVVTVMATPGFVFEFSRORGLADGRCKSPAAADGTGWE 2020
 QY 723 GVGMLLVERLSDAVRLGHRVLAVVRSANVODASNGLTAPNGPAQERVIRQALANAGLS 782
 DB 2021 GAGLVLLERISDARRNGHVLAVRGSANVODASNGLTAPNGPSQQRVITQALASGLS 2080
 QY 783 VADVVEHGCTGTTGCDPIEAQALLATTGQ-RAGRPMLLSGSKNIGHTMAAAGVGV 841
 DB 2081 VSDVDVAEAGCTGRIGDPIEAQALLATTGOGSDSRPMLSGVSKNIGHTTAAAGVAGV 2140
 QY 842 IKVMALRBSGVLPRTLHVDPSPQVMSAGVRLTEAVPWPDAAGRLRRAGVSSFGIG 901
 DB 2141 IKVMALRBSGVLPRTLHVDPSPQVMSAGVRLTEAVPWPDAAGRLRRAGVSSFGIS 2198
 QY 902 GTNAHVILEEAPAAAGCVAGGVLEGAPGLATSVASVAPVAVASVAVESVPPVPPV 961
 DB 2199 GTNAHVILEEAG-----SKTPBETADKSGPDSST-VQDPA-V 2233
 QY 962 PVFVSARSEAGLRAQAEALRQYVAVRPDVSADVGAGLACGRVLEHRAVVLAADEBELV 1021
 DB 2234 PLIVSGRTPAALSAQASALSLYIGERGDIITLDAASFSLASSRAALEERAVVLGADERETL 2293

Qy	1022	QGLBAAAGBERRRHTTTHAAGSGBRGVAVRPGGSGGGMAGBGRVLLASSVPVARRMOC	1081
Db	2294	SGLELBSGREGASGVSSPVS---GGGVFPAGGGGGLMGKGLSVFPEFA---DAF	2347
Qy	1082	BEALAP-----WDMSVVDILRRDAGDAVWEERADVQVLFESVWVSLAALRMSYGBERDA	1136
Db	2348	DEACSGIDAHIGORGVADVFGSGDG-SILDRITLMAOGSLFALQUGLLSLGSSVGRGV	2406
Qy	1137	VLGHSQGEIAAAHVCGLSLKODAKTVALRSRDLAALVGRGMAVPLRAQVEBOLGER	1196
Db	2407	VLGHSVGEFAAVALVAGVLSLPRDAARMVGRARLMQALPRSGAMLAVALAGEQLRPLLDR	2466
Qy	1197	MAGLWMAVAVGRPSTANVSGDAEAVDLVACAGGVAVARILPRDUASHCHNQRVREEL	1256
Db	2467	VDG-AGIAAVAPESVVLVSGREVLDDTAGALDGGALIMRLRVSASHASTYRMDMIOEF	2525
Qy	1257	LELLGDSPQSPGVPRFSTVEGTWLDITTLDA--UYWRNLHQVRPFSDAVQALADHGR	1314
Db	2526	AEIARSDVYRRGDLPRVSTLGE-LDTAGWATPEWYWRQVREPRFRFDGVRLAQGSA	2584
Qy	1315	VFVEVSRPRLVPRAIEDTETDAEDVTAIGSLRGSDNTRRPLRALANTHTTGSTPTW	1374
Db	2585	TIPEISRPATLSALPRD-CHSMAOCAMRIPMLRKQRTETETVVAVALAAHNRGV--PVEW	2641
Qy	1375	HHHTHHHTHHRHHTLPTVPRQOHUWYLESSQPGSSG-----	1417
Db	2642	SAVFRAGTAR----RVELPTYARQORUYMETSDYGVGTIGLAAHPRLLGAVVALAD	2686
Qy	1418	-----AGAGSG-----	1423
Db	2697	GDGMVLTSRGLSVGTHPIRLQHRVLGEVUVPRGALIMMLHAGARLCSQVBELETPELV	2766
Qy	1424	-----AGSGRAGTAGG-----TVEVERFYMDAVAROLE-----	1452
Db	2757	VPRBAAGSGRPGAGTGVSIETAEERVKTDAIEIQLLVNAPREGGRVAVSLYSPAG	2816
Qy	1453	-----TVATT-LAVP-----PSAGLPTVPRAL-----S	1474
Db	2817	SRGGSWTRHATGELVGVTTGGBAVFDMSAEGESIALDEPVYALLANGSFEYRLFOGLQ	2876
Qy	1475	AMNRHONDOARI-----NTW-----	1489
Db	2877	AMRGDEVLAEIARABADAMASGLDRPALLDQLSALSDRREGGAWLPRSGTVE	2936
Qy	1490	-----TY	1491
Db	2937	LSARAGTISRVLERRRDAISVAVMDESGRLASIDSLRYKSVSSQOLANRDAVRDALF	2996
Qy	1492	QETKRLRLPTTHORNOYMLAIRETOHNNHNTILNNHNN--GIPIRLNHTNTN	1549
Db	2997	EUTHERVAVQST-EEGR-MAL-IGOTASCODLKLADDSADRCADLALAEKDDSSALV	3053
Qy	1550	POHLNHTNHTROQAKHNTGTGATLSLALDETNRHNNHNTPTGTLNLTQTNHOT	1609
Db	3054	PDVV---UYCAGEQADREGGA-----ALAEQOQ-----TLALQOMLAEPRL-	3093
Qy	1610	HRPRPLMJATTNATTNTRNDRLPTHTQACTGALATTLLENPTHTAGIIDLEPTT-----P	1664
Db	3094	AEARLVVUTCAAVUTAPASDGAISELANALYGLGLRAAOVENPQGFV-LADVDGAEASWRA	3151
Qy	1665	TPHTLOHLTQTLGOHNOTOAIPTGTGHTYRLPTPTLTPHQRPTPRPHGTLITGGTG	1724
Db	3152	LPRLAGSN-----EROLAKRGAVARPLASVAGQIDVRAVADPRKTVLSGGGTG	3202
Qy	1725	ALATHTLNTLTTHORONLLTSRTGRTPHNAONLTLOOKSGLNLTTCDSNSPDLO	1784
Db	3203	LLGGAVAHNLVTGERVRLVLIGRGMDARGTIVELVGLNGAGAVVUDVADVRADLVE	3262
Qy	1785	OLANTIPRONLTUYIHTAGLIDATLNLPTPTOLANVLKAKASHNLHOLTONTPLTA	1844
Db	3263	SLLAAPVREFPLCGVVAAGALADVITSLSPDDGAVFGPRKAGAMWMLHEITRDTLSF	3322

Oy	1845	FVYVSSAAATFGPCQGANVYAAANVADALAHHRHTHLLPSTSIAMGW--ONGSLADBDKA	1903
Db	3323	PALFSSISGVAGAPGQGNVYAAANFIDLALHYRSGSLPAVSLAMGLMEOPSGTET---	3379
Oy	1904	RAYLDREGRPMSEBELT---AAVTOAIADTERPVYIADIWMSKIHTSOTSDLSAAR	1960
Db	3380	LSEVDRIARIANPDLSTKEGRLFDQGLALDRAAVPALDRIPLFLEQARSSL-----	3434
Oy	1961	EREPAVORPTPELHKTLLAHQTSADORAALLBELVRDVAVLRHADPKAIAPDOGRFAL	2020
Db	3435	---PALTLTALVPPIRRNRRASTELADEGTLGVREHEAAVAVLGSSAAVGVERRFRL	3491
Oy	2021	GFPSLTAVEERNLIKATGRLRPSVLYFHDPTPAKLVHLONGLRGTAASAPSAAV-T	2079
Db	3492	GFPSLSEVEELNRNLAGVLRPLPAVAFDPTPALRAFHLQEL---ADELATTPAVTT	3548
Oy	2080	AEASVTPP--IAIVGMACRPPGQTSADDPMDLISSEQDAIGCFPTDRCMDLTLYPDPP	2137
Db	3549	TRAPVAADDLVAIVGMCGRFPQVSSPEBELMRLVAGVDAVADPPADRGMDLGLFPDP	3608
Oy	2138	DHPGTCYTRNGGLYDAGHDAEPFGISPREALAMPQORLLLETAMETTEHAGINPHTL	2197
Db	3609	ERAGKTVYRGCAGFLTDDREFDAGFGCISPREALAMPQORLLLETLSWEALERAGIDGSL	3668
Oy	2198	HGFTPTGPTGNODVALR--VHNAQSTDBERPALTGTAHSVIGSIISTPGEGEPAVVD	2256
Db	3669	RGSSTGVFAGIMYHDYCARPASRPPEFEGYLGNGSGVASGCIAYSFGEGGAVVDT	3728
Oy	2257	ACSSSLVALHLACOLAGAGECSMALAGVYVMSPGAFVFFSRORGALADHCAFPALAA	2316
Db	3729	ACSSSLVALHLLAGOSLRSGCDLALAGVYVMSPTGFVEFSRORGALADGRCSFABSA	3788
Oy	2317	DGTGMEGEGVMLVERLSDAHNRGHRVLAVVRGSAVNQDASNGLTAPNGPSQQRVTRQA	2376
Db	3789	DGTGMEGEGAGVLLERLSDARRNGHRVLAVVRGSAVNQDASNGLTAPNGPSQQRVITQQA	3848
Oy	2377	LAANGLSAGVDVAEAGTGTTLGDPRIEADQALLATYGQDPAEGBRPLGSKVSNVHTQA	2436
Db	3849	LASGLSVSVDVAEAGTGTTRLGDPRIEADQALLATYGRDRPCGRPLGSKVKSIGHTQA	3908
Oy	2437	AAGVAGVYKVMVMLRHGLLPRTLVHDPSPHHVMSAGAYOLLTEVPMPEGEGRLRAGV	2496
Db	3909	AAGVAGVYKVMVMMRHOQLERTLHVDPSSQVDMSAGRVOLLTEBNTWPR--DSGRPCRVGV	3967
Oy	2497	SSFPGVSGTNHVLLEBARADVPBGCPAPGEGD--AGSDDEAAAGSPGVWPMVLVSAKSOPAL	2555
Db	3968	SSFPGISGTMHVLLEGS-----TGQMDQAAERDSSPVLDVPVPMVSGKTEBAL	4017
Oy	2556	RAQQAOLHMLTDHPRGLDLADVYTLAHARAVDPHRAITLIAADBDTGLDALQALAGEPH	2615
Db	4018	SQAQAATATATYLDQVNDVSPLDVGISLAVTSSALDERVAVVGSDBDTLISGLNALAAG---	4074
Oy	2616	PAVTHSSAPGSGTGEAAGKTAFCGSGQSGORGBMAGLVNTHRVFAALINDICTHLDPH	2675
Db	4075	---HEAAGVYTGVEVGIGRTGTFPAQGGGQGLCMGKRLLSRRFAFGARDEACAEIDAN	4130
Oy	2676	LDRP--LLPRLTONDNEDNEAALLQOTRPAORALFAFOVALNRLITDGYNHTPHVYAGH	2733
Db	4131	LGRVGRVADVFGSDE-----SLDRITLMAQSLRALQGLWELL--GTWGVRSSVVLGH	4183
Oy	2734	SLGEITLAHLAIGIITLTDAITTLTORATLMQTMPP--GTMTTLANTPHNITNHLTAHENDL	2792
Db	4184	SVGELAAAFAGAVLSMAEARLAVGARARLQALPFGSGAMLAVSATBARVGPLDGVLRDV	4233
Oy	2793	ALAIINPTSLVUSGTHVQHTITTLQOOGIKTKTLPTNNAHPSHTNPRLYNLQHTQ	2852
Db	4244	GVAAVNAPGSVUSGDRVDVLDGLAGRLDGGGHSRKLAVSHAFRSHMDRPLAFELAR	4303
Oy	2853	TLATVPHRPTLTANT---PPDQLTPNHWTOOARNTVUYATTTQTLQONGVITYELG	2908
Db	4304	SVUDRSRPLRPIVSLTGNLDVUGVMAPIREVIWVQVRVAPRADGVQALVDOGVDTIVELG	4363
Oy	2909	PDNLTTLTLTHNHLNPRPTT--LITLPHNHNROTLLTNLAKTTTHNPHNHTHNDOPHT	2967

```

Db 4364 PDGALSSIVGECVAESGRATGIPIVRRDRDEVRTYLDALAQ-----HT 4407
2968 H-----THLDIPYFQHHNHWLESTQPGANNVSAAGIDPTEHPLGATL 3012
4408 RGAVMDGSGFFAGTRATQVDLPYAFQFORWLEPSD--SGDVTGVLGAHEHPLAVV 4465
3013 ELATDGGALLAGRLSLNSHPWLADHAGVGVTLISGATELELALAGTVGCDRDETLH 3072
4466 PVAGGDEVLTLGRSLVGHFWLAEHRLVGEVVGTLALLENAMKASGVGGEREELTLE 4525
3073 APVLPVDDGVSVGVVGAADGEGRLVSVYARSGSAGCGGAGVWTCASGVLYEAA 3132
4526 APVLPVPERGAAYOLAVGAPDEARRSRISQIYSRGAD-----EDGWRRIASGLLQAN 4578
3133 AGGVVVGGLGVPRGAVNVDVGVDRILAGACVIGPVFSGLRVWVDDGDLILAECVL 3192
4579 A---VPPADSTAMPDPDAGQVDLAEFYERLEMERGLTYGPVFOGLRAAMRHGDDI-FAELIA- 4634
3193 PEEAMGDPAAGFGLHPALIDGVVQPLSVLLPGTGFGEAGAGFGEVPRVAVWGVSLHRAG 3252
4635 ---GSPDASGFGIHPALIDALHLMAL-----GASPDSEARLPIFSMRGAQLYRAE 4681
3253 VTGVRVSAVAGRGGRBRASVVVGVDEACVAVSDRLRLRPVDMGQLRAVSVSAGRRGS 3312
4682 GAALRVRLSPISGSG---AVSLTLVDATGRVAAVESLSTRPVSTDOIGA---GRGQER 4734
3313 LVANOW---AEVGPVPOCGQMA-----MHEDVGESG-----GGPV 3345
4735 LHAWEWRBSASASAMSLTSCAVVGLGEBPEHAAKTKTQVESHADLASLATEVAKRGA 4794
3346 PGVVVLCRDPAGAGGGGGGGGGGVGEV-----VGVGLGVVOGWLGERFASRLVVV 3398
4795 PGATVIVPCPRRA-----MQLPTAARRATQGAAMMLQOMLADRFVSTRILL 4843
3399 TRGAVVAGPEDDGPVDDVVGASVWGLVRSQAQEHPRFVLLDIDTGTDLDTGAGAGVND 3458
4844 THRVAIVSAGEDVDLVHAPLMGLVRSQAQEHPRFALIDMDDERASQT----- 4892
3459 GGRVAAVAGGEPRGLAVRGERLLAARLKLESSDDVAORSGDPRARRSDVPARSGGPV 3518
4893 --ALAEALITAGEAQLAVRSGVLAIPRIGQ----- 4919
3519 ARRSVDVSGREVLTPMLSGSVLTGTGVLGAAVARHLAGVGVRRDLILVSRRPDAPGA 3578
4920 ---VKVSGGEAFRW-DEGTVLVGTGGGLGALLARHLVSHGVRRHLLLSRRGLAHPGA 4974
3579 EGLRAELALGAERITVACDVGERRERVLIEGVPAQCPLTGVVHAAVGLDDATIASLTP 3638
4975 DELVAELEQAGADVAVVACSDRDSLARLVASVPAERPLRVVHAAVGLDDGVLMSMSP 5034
3639 ERIGTVPAKVDAALLDELIRGMELSAFVLFSSAGILGSAGGQNVAAANALDALAYR 3698
5035 ERLDAVLRPKVDAMWYHLELRELGLSAFVLFSSVAGLFGAGGQSNVPAAGAFALDALHC 5094
3699 RRAAGLGVSLAMGLMEASGMTGLAGTDHRRIRISGLHMSPTPDALALFDALALDRP 3758
5095 RQAGGLALSLASGLMASIDGMAGDLAADVRLSRAGIGPLSAPGALFDDAVGDEP 5154
3759 VLPADL-----RPADPLPDLLODLIPATRRRTTTTGGADNGAQHARLAGQ 3808
5155 LLAIVRLDVEALRVOARSVQTRIPEMLHGMAHGSPRRTPFSRV-----EPLHERLAGL 5208
3809 THEQHTLLALVNSHATVYGHPTPTIPDRAFRLGPGSLFVAVELRLRLSTTGLRL 3868
5209 SEGRROQVLORVADIAVVLGHGRSSDVIDEKLPLAELGFDSLTAIBLRRLAATYGLRL 5268
3869 PTLAAPHNPPTTLTHLHTQL---OPOP-----DNAVAVLAEKLBSALSALDKTD 3919
5269 PATLAFPHGTRAAALAGVCAQLGTRATAPARPRTDNDATBEVRLPQQAAYAAAGILDMGD 5328
3920 SASERVTLR 3928

```

```

Db 5329 LVKVAAGLR 5337
RESULT 11
ID 030764 PRELIMINARY; PRT; 4340 AA.
AC 030764;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Polyketide synthase modules 1 and 2.
GN NIDAL.
OS Streptomyces caelestis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriales;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
CX NCBI_TaxId=36816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL-2821;
RX MEDLINE=98053867; PubMed=9393718;
RA Kakavas S.J., Katz L., Stassi D.;
RT "Identification and characterization of the niddamycin polyketide
RT synthase genes from Streptomyces caelestis.";
RL J. Bacteriol. 179:7515-7522 (1997).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL; AF016585; AAC46024.1; -.
DR InterPro: IPR002106; AAcRNA_ligase11.
DR InterPro: IPR002127; Ac transferase.
DR InterPro: IPR002198; ADH short.
DR InterPro: IPR004410; FAD.
DR InterPro: IPR000794; ketoacyl-synt.
DR InterPro: IPR003880; Plnane attach.
DR Pfam; PR00698; Acyl transf; 3.
DR Pfam; PF00106; adh_short; 2.
DR Pfam; PF00109; ketoacyl-synt; 3.
DR Pfam; PF02801; ketoacyl-synt_C; 3.
DR Pfam; PF00550; PP-binding; 3.
DR TIGRFAMs; TIGR00128; fadp; 3.
DR PROSITE; PS00339; AA TRNA LIGASE_II_2; UNKNOWN_1.
DR PROSITE; PS00075; ACP DOMAIN; 3.
DR PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
KM Oxidoreductase; Phosphopantetheine; Transferase.
SQ SEQUENCE 4340 AA; 457589 MW; D59A734C5F8795D CRC64;

Query Match 34.8%; Score 7244; DB 2; Length 4340;
Best Local Similarity 42.6%; Pred. No. 9,1e-303;
Matches 1764; Conservative 421; Mismatches 1438; Indels 520; Gaps 99;

7 GEEPRPAGVGLVADGAGVVFPPGQGPQMPGKGRRLD-----ASDVFRESVRA 59
518 GQTFR---GVRISTD-ADGRLLALFTGGAGHPGNGELIYTDHPHAAALDEVCELR 573
60 C-----EAAFAFYDMSVEQVLRLSDPADGLDRDVVQPTLEFVNI SLAALWRSQVPE 113
574 CGTQRLREWVFTD-----DQPDL-IDRTEYDPAFLQATLVRTLTATAGTQA 620
114 CAVLHSLGRIIAAHVSGGLSLADARVVTLSQAQTLTLAGTALVSVATPELLPRLA 173
621 HVLVGHVSVEITTAHLAGVLDLPDAARLITARAHVWGOLPHGAMLSVQAHEHL----- 675
174 PWTEDNPAR---LAVAAVNGRSTVSGAREVADVLDLTAQVTRMIPVDPVAPASPL 230
676 ---DQLATHGVEIDAANGPTHCVLSGPRTALEETRAQHLREONVHTWKSHASHAL 731
231 MYAIEERVSGLPITPRSRIPFSSVTCGRDLTREL-DAAYVNMSTVAFEPAPRL 289
732 MDPMLAGFRNTLNTLVNQPTTIPILISNLGQALDPHMLCPDVIDHABHTVFAAVOT 791
290 LLOQGPTEVMSPHVLTNGL-----QELAPLDGDTTGRA 325
792 AHQGTYYVLEIGPHFTLTLLHHTLDNPTTITLHRRRPERPETLQAIAAVGRVDGID 851

```

QY 326 DTVMGLTRBQOGL-----DHFLTSLAQLRGHGETSATVLSARLTALSPTQOOSL 377
 Db 852 MAVLCSGRPRRVELPTVAFQRRTHMA BGLTP--NHAPDRPAEPQRAAVGVSRXA- 908
 QY 378 LLDLVRAHTMAVNDGONERTASDAGPSAGFAHGFPSVWGVELRNLSKATGLRLVTL 437
 Db 909 LVRVVGSTTASVGLDGPDEVALD-----RPTSGQLSMTAVELAGLGLTAAGVALDPTL 964
 QY 438 IFDHTPAVAARLRTAALGHLDDEDTA VPDPSPSGHG--TAAADPLIATIGMACRPPG 495
 Db 965 VTELPTRAVAADHIAKTLG---ESAADDOGVNKGTEBAKADPLAVIGICRPPG 1021
 QY 496 VRSEKLMELASGDAIGEPPTDRGWPTEQRHQAQDPTQPTFYPOGGLHDAHNDAG 555
 Db 1022 VATPDDLMELVASGDAISTFTDRGMDLGLYDPDSTPEKSVYRRGFLHDAQDAE 1081
 QY 556 PFGISPEBALAMPQORLLLETSEWEAPERAGIDPLSVRSGRTGVFAGALSFDYGPMDTA 615
 Db 1082 PFGISPEBALAMPQORLLLETSEWEAPERAGIDPLSVRSGRTGVFAGALSFDYGPMDTA 1138
 QY 616 S8SEGAADVEGHILGTGSVLSGRVAYSFGLEGPAITVDGCSASLTTLHACOSLASGE 675
 Db 1139 -H8TGDYEGFLTGTTASVAGRIATVLTGRPALTVDPACSSSLVALHLAVOSLRGE 1197
 QY 676 CTLLAGVSVMTLGNFIEFSRQGLSVDRCKAVSAADGTGMEGVGLIVERLSDA 735
 Db 1198 CDALAGCTVMSGKMFVEFSRQGLAPGRCAPADADGTMAAGVGLIVERLSDA 1257
 QY 736 VRLGHRVLAVRGSAVNDGASNGLTAPNGBAQERVIRQALANAGLSVADVVEGHGTG 795
 Db 1258 ERLGHRVLAVRGSAVNDGASNGLTAPNGBAQERVIRQALANAGLSVADVVEGHGTG 1317
 QY 796 TTLGDPTEAQAALLATYGORAGDRPLMLGSLKSNIGHTMAAAGVGVKMMVALBESGLPR 855
 Db 1318 TALGDPTEAQAALLATYGORAGDRPLMLGSLKSNIGHTMAAAGVGVKMMVALBESGLPR 1377
 QY 856 TLHVDPSPQVMDSAGAVRLTTEAVPWPBGDAAGRLRAGVSSFGIGTMAHVLIEEPAA 915
 Db 1378 TLHVDPSPQVMDSAGAVRLTTEAVPWPBGDAAGRLRAGVSSFGIGTMAHVLIEEPAA 1435
 QY 916 GGCYAGGVLEBAGFALISVAESVAAPVAVSAPVAVSAPVAVSAPVAVSAPVAVSAPVAVS 975
 Db 1436 GGCYAGGVLEBAGFALISVAESVAAPVAVSAPVAVSAPVAVSAPVAVSAPVAVSAPVAVS 1464
 QY 976 QAEALRQVAVRPRVSLADVAGIACGRAVLEHRAVVLADRELVGIGALAA-----GE 1031
 Db 1465 QAEALRQVAVRPRVSLADVAGIACGRAVLEHRAVVLADRELVGIGALAA-----GE 1523
 QY 1032 PDRVTTGHA PGCDRGVVFVFPQCGQMGAGVRLASSPVFARM--QACEALAPMD 1090
 Db 1524 PDRVTTGHA PGCDRGVVFVFPQCGQMGAGVRLASSPVFARM--QACEALAPMD 1569
 QY 1091 MSVVLDLRDAGAVMERADVOPVFSVMSLAALRSYGIEBDAVLGHSOGEIAAAV 1150
 Db 1570 MSVVLDLRDAGAVMERADVOPVFSVMSLAALRSYGIEBDAVLGHSOGEIAAAV 1629
 QY 1151 CGALSLDAKTAVALBRALAAVRGSGMASVPLRABEVQOLIGERAGLMAAANVGR 1210
 Db 1630 CGALSLDAKTAVALBRALAAVRGSGMASVPLRABEVQOLIGERAGLMAAANVGR 1686
 QY 1211 STAVSGDAEAVDEVLAYCAGTVARARI PVYASHCPHVOPLREILELLELSDISPOSGV 1270
 Db 1687 STAVSGDAEAVDEVLAYCAGTVARARI PVYASHCPHVOPLREILELLELSDISPOSGV 1746
 QY 1271 PPSSTVEGMLDPTTL--DAAYVRNHLQPVRFSDAVOALADDGHRVVEVVSPPHTVPAI 1329
 Db 1747 PPSSTVEGMLDPTTL--DAAYVRNHLQPVRFSDAVOALADDGHRVVEVVSPPHTVPAI 1806
 QY 1330 EDTTEDBAEDVTAIGSLRGNDTRRLTALAHTHTTGIGTPTMHHNHTHHNHTPHRT 1389
 Db 1807 EDTTEDBAEDVTAIGSLRGNDTRRLTALAHTHTTGIGTPTMHHNHTHHNHTPHRT 1853

QY 1390 HLDLPTVPEQHONYME-----SSQPGAGSSGAGAGSAGSAGTAGTAEVESRP 1442
 Db 1854 HLDLPTVPEQHONYME-----SSQPGAGSSGAGAGSAGSAGTAGTAEVESRP 1888
 QY 1443 ---WDAVARODETVAITLAVPPSAGIDTVPLLSAMHNRHQDAINTTVOETKPLT 1499
 Db 1889 WDAVARODETVAITLAVPPSAGIDTVPLLSAMHNRHQDAINTTVOETKPLT 1909
 QY 1500 LPTTHQPHQWTLAIPEOTGNHPHINILINLHNGITPRLPLNHTHTNQHLNHTLHN 1559
 Db 1910 LPTTHQPHQWTLAIPEOTGNHPHINILINLHNGITPRLPLNHTHTNQHLNHTLHN 1934
 QY 1560 TROQAQNTGATIGLSSLADETP--PHNHPRTGTLN-----LTLQTNQTH 1610
 Db 1935 TROQAQNTGATIGLSSLADETP--PHNHPRTGTLN-----LTLQTNQTH 1990
 QY 1611 PPTPLVATATNTATTNPNBLTPT--OAGTWGLATTLLENHTHTAGIIDLTPPTPT 1668
 Db 1991 PPTPLVATATNTATTNPNBLTPT--OAGTWGLATTLLENHTHTAGIIDLTPPTPT 2047
 QY 1669 LOHLOTLOPHNQTOAIRTTGTHTRRLPTTLTLPHTPPTPTPTPTPTPTPTPTPTPTPTPT 1728
 Db 2048 LOHLOTLOPHNQTOAIRTTGTHTRRLPTTLTLPHTPPTPTPTPTPTPTPTPTPTPTPTPT 2105
 QY 1729 HLTHNLTHQPTQHLNLSRTGPRTPHQAHLTTLOOKGILHTTTCTSNPDLOQLN 1788
 Db 2106 HLTHNLTHQPTQHLNLSRTGPRTPHQAHLTTLOOKGILHTTTCTSNPDLOQLN 2162
 QY 1789 TTPQHPLTVITHTAGIIDLDTATLTNLTPTQANVLAHAKHSAHLHOLTHTPLTAVLY 1848
 Db 2163 TTPQHPLTVITHTAGIIDLDTATLTNLTPTQANVLAHAKHSAHLHOLTHTPLTAVLY 2222
 QY 1849 S8AATFAGPQOANVAAVADALAHNHTHNLPTSTIANGWQONGSLADSKAAYLD 1908
 Db 2223 S8AATFAGPQOANVAAVADALAHNHTHNLPTSTIANGWQONGSLADSKAAYLD 2281
 QY 1909 RRGFRPMSBELATAVTOAIADTERPVYADIDMSKIEHTSQTSDVLSAAREPAPVR 1968
 Db 2282 RRGFRPMSBELATAVTOAIADTERPVYADIDMSKIEHTSQTSDVLSAAREPAPVR 2340
 QY 1969 PT-----PPALHKTLANQTSADQR--AALLEVRDVAVALRHADPKAIPQSPFALG 2021
 Db 2341 PT-----PPALHKTLANQTSADQR--AALLEVRDVAVALRHADPKAIPQSPFALG 2400
 QY 2022 FDSLTAVERPNLLIKXTGLRLPSIVEDHPTPAKLVHLOLONLGR-----AASAPSAAY 2078
 Db 2401 FDSLTAVERPNLLIKXTGLRLPSIVEDHPTPAKLVHLOLONLGR-----AASAPSAAY 2460
 QY 2079 TAEASVTEPIAVGMA CREPGVTSADDFMDLISSEQDAIGCFPTDRGMDLDTLYDPDP 2138
 Db 2461 TAEASVTEPIAVGMA CREPGVTSADDFMDLISSEQDAIGCFPTDRGMDLDTLYDPDP 2519
 QY 2139 HPGTCYTRNGGLYDAGHFDGAEFFGISPREBALAMPQORLLLETMAWETIEHAGINPHTL 2198
 Db 2520 HPGTCYTRNGGLYDAGHFDGAEFFGISPREBALAMPQORLLLETMAWETIEHAGINPHTL 2579
 QY 2199 GTFPGVETGNQDVALRVHNGQSTDFALCTASVSGRISYTFGFGPAGVAVDTAC 2258
 Db 2580 GTFPGVETGNQDVALRVHNGQSTDFALCTASVSGRISYTFGFGPAGVAVDTAC 2639
 QY 2259 SSSLVALHLACQALRAGCESMAGAVTMSFGAFVEFSRQGLAADGCHKFASAAADG 2318
 Db 2640 SSSLVALHLACQALRAGCESMAGAVTMSFGAFVEFSRQGLAADGCHKFASAAADG 2699
 QY 2319 TGMGEGVGLVVERLSDAHNRHVRVLA VVRGSAVNDGASNGLTAPNGSPQORVIFQALA 2378
 Db 2700 TGMGEGVGLVVERLSDAHNRHVRVLA VVRGSAVNDGASNGLTAPNGSPQORVIFQALA 2759
 QY 2379 NGLSAGVDVAEABHGTGTTGDTTEAQAALLATYGODRABGEGVGLKSVKSNHGHQAAA 2438
 Db 2760 NGLSAGVDVAEABHGTGTTGDTTEAQAALLATYGODRABGEGVGLKSVKSNHGHQAAA 2918
 QY 2439 GVAGVVKMMVALRHGILPRTLHVDEPSPHVDSAGAVQULTETVWPFGEGRLRAGVSS 2498

```

Db 2819 GAGTIGGIMHNGTLPTLTHVDAPTSQVDMEAGSLQLLEAPRWPADPRPRRAGISA 2878
Qy 2499 FGVSGTAATVILEAPADVDVGGPPAGEGAGSDEDAAGSPGVWPLVSAKSPALRAQ 2558
Db 2879 FGVSGTAATVILEAPADVDVGGPPAGEGAGSDEDAAGSPGVWPLVSAKSPALRAQ 2925
Qy 2559 AQAHAHLDHPGDLADVGCTLAHAAPFDRATLLAADR---DTFQALQAL-ANGER 2644
Db 2926 AARLDSTL-HHTQADPDIDIAHATCTCHFKRAVITGRRTTEHKLHLDALQSTANR 2984
Qy 2615 HPAVHSSAPGCTGGAAGKTAFCGSGGTORPGMAHGLYHTRVPAALNDICTHDP 2674
Db 2985 HP-----RLTLFTGGQGHQMGQELYTTPDHPFAALDEICELO- 3025
Qy 2675 HLDHPLPLTQNDND---NEDAAALQOTRYAPALFAPVALHRLTDCYHITTPRYA 2731
Db 3026 -----RCGTQNLREWFEPDQDLDLREYTOPALFALQALYRLLT-ANGQAHVL 3077
Qy 2732 GHSIGEITAAHLAGILTLTDATTLITQATLMQMP-PGTWTLHTTPHHITLHAEN 2790
Db 3078 GHSIGEITAAHLAGILTLTDATTLITQATLMQMP-PGTWTLHTTPHHITLHAEN 3135
Qy 2791 DLAAIAINTPTSLVIGS---TPHTVOHITTLCOQOGIKTKTLPTHNAFHSPTNPILNQ 2846
Db 3136 GVEIAAVNGPHTCVLSGPRALBESTAOHL---REQVNRHTWLKVSHAFSALMDPMLGA 3191
Qy 2847 LHQGTQTLTHPHPTPLTANT---PPDOLLTHYHTQOARNTVDATTTQTLHQGVY 2902
Db 3192 FRDLTANLTYOQPTPLISNLTGQIADPNHLCSTPDYIDAHKHTVRADAVQTHADRRT 3251
Qy 2903 TYELGPDNLTTLTHNLPRPTTLTLTHPHNPOTHLTLNLAKTTHPHNHYTHD 2962
Db 3252 TYELGPDNLTTLTHNLPRPTTLTLTHPHNPOTHLTLNLAKTTHPHNHYTHD 3302
Qy 2963 NQPH-TH---THLDPTTFQHHNYWLESTPGANVSA-----AGLDPTENPLLGATLE 3013
Db 3303 TTPHPSHLPAGVSELPAVYFQRAVYMMNSAHIGRSDAEATRLGLAQOHPLTGTATS 3362
Qy 3014 LATGAGLLAGRLSRHPLADHAGVSTVLSCATFELALAHGYVCDRVDLTHA 3073
Db 3363 IVSGGATLLTGRVSLAHPMLADHTVAGAVLPGTAADLLHNAEETSGSTVELTHV 3422
Qy 3074 PLVAVPVGAGVVOUGVAAADGGRRLVSVVARGSGAGSGGWTCHASGVLEMAA 3133
Db 3423 PLVAVPVGAGVVOUGVAAADGGRRLVSVVARGSGAGSGGWTCHASGVLEMAA 3476
Qy 3134 GGUVV---DGLAGVPRGAVAVDVGVRLAGAGCVLGPFGSLRAVWRDGDLLAEV 3190
Db 3477 AAPAVDADMAAGAMPFGAEAVVDELTAFAAAGYDYGPAFTGLGAMRGEVPAV 3536
Qy 3191 CLPREANG---DAAGFGLHPLLDGVQPLSLPGG--TGFGGAGRGSEVTRPAVWG 3245
Db 3537 RLPRVAVSGEERERFGVHPLFLDALHPRV---GGWOBESGDADGGSTLLPFGWOG 3593
Qy 3246 VSLRAGVTGVRVSVAVRGSGREAVSVVVGDEAGVPVAVDLELTPV----- 3295
Db 3594 LTLVGTAEQVRVVL-APSDGG---AFSVRAADPAGHVLADLVLRVSVSTPGALPA 3649
Qy 3296 -----DMGOLRAVSVSAGRRGSLYAVQMAEVGPVCGQAM-----AMHEDVESGGG 3343
Db 3650 APWHIMQPVDSISLVDTRR---YAVVGADETAELVRAFLPDAHTTAADLALRAA 3706
Qy 3344 -----PVGCVVTLRCPDAGAGGGGGGGGGGVEVVGVLGVYQGLGLER-ASRL 3395
Db 3707 PDAAVAPTPGALVVALIPGSPPARPEATAPPRATLHGLSLVQALIEBERLTTGTL 3766
Qy 3396 VVYTRGAVVAGREDPDVGVASWGLVRSQAQEHPRFVLLDPTGTDLTDGAGAGW 3455
Db 3767 AVVTRGAVVATQTVTP-DLAGAALMGLVRSQSTPGFRLLDVA-----DSPESSRA-- 3818
Qy 3456 GVDDGRVAAVAVACGEPOLAVGERLLAARLKRLESSGDPVQORSGDTRARRSDVPAORSG 3515

```

```

Db 3819 -----LATALADEPOLARAGVMLAPAL-----TPAEAAG-----SEAPAD--- 3855
Qy 3516 GVPARSVDSGSEVLEPMLSGSVLTGCGVLAAGVAAVHRLAAGVGRDILLVSRGPD 3575
Db 3856 AAPTAHGMND-----GTVLLTGCTGALGRRVAHRLAHGVRRLVSRSGENA 3903
Qy 3576 PGAEGRABEALAGAEVRIACVGERREVRRLLEGVPACPLTGVVNAAGVDDATIAS 3635
Db 3904 PESALERELTELGVHTSFVSCDLADPADVRAKAVAGVPSDHLPLGVHTAGVDDGAL 3963
Qy 3636 LTPERLGTVPAAKVDALLIDELTRGMELSAFVLESAGTICGAGCGYAAANAALDAL 3695
Db 3964 LTPERIDTVLRAKADVRHHEATLGHPLAFVFGAAAGLRLPGQASVYAAANAALDAF 4023
Qy 3696 AYRRRAAGDPVSLANGMEASGTHAGTDRHSILRSGHPSGPDALFPAAL-- 3753
Db 4024 ALRRAEGLPAVSLANGMEASGTHAGTDRHSILRSGHPSGPDALFPAAL-- 4083
Qy 3754 -----ALDRPVLPAD-LR-----PAPPLPDLDPATRRRTTTTGGAD-NGA 3799
Db 4084 YRDGPALVPLLDSTALRTAGENGPA-SVPLLSLVPSHRRRTTATSVRDESGP 4142
Qy 3800 QHARLAGQTHEQHTTLALVRSHTATVGTTPPTTPDRAFRDLGFSVLAVERLR 3859
Db 4143 GPGARLAAAPAKSRMTVLLDTVSEOVAGVGHASAAEIDBERPFHIGPDSLATLELR 4202
Qy 3860 LSRRTGLRLPTLAFPHNPPTTLTHNLHQLQOPDNVA-----PYLAELDKLESALSA 3914
Db 4203 LGRVGRRLPSTLAFPHNPPTTLTHNLHQLQOPDNVA-----PYLAELDKLESALSA 4261
Qy 3915 LDKTDSASERVTLRKLSMLRWNAPOHTAESADD-----EKTSATEAEIRFKFD 3966
Db 4262 MDMEDGRADVQRVGLLTLALDRP-----ASGADSTRAAATVADRLEATDDEIFAFLD 4317
Qy 3967 NDL 3969
Db 4318 EOL 4320

RESULT 12
QX993 ID QX993 PRELIMINARY; PRT; 6797 AA.
AC QX993;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Polypeptide synthase (PIM1 protein).
GN PIM1.
OS Streptomyces natalensis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=68242;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9214571; PubMed=10187796;
RA Aparicio J.F., Molina A.J., Ceballos E., Martin J.F.;
RT "The biosynthetic gene cluster for the 26-membered ring polyene
RT macrolide pimaricin. A new polyketide synthase organization encoded by
RT two subclusters separated by functionalization genes."
RL J. Biol. Chem. 274:10133-10139 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Aparicio J.F., Fouce R., Mendes M.V., Olivera N.;
RT "A complex multicenzyme system encoded by five polyketide synthase
RT genes is involved in the biosynthesis of the 26-membered polyene
RT macrolide pimaricin in Streptomyces natalensis."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ132222; CAB41041.1; -.
DR EMBL: AJ178573; CAC20931.1; -.
DR HSSP: P25715; IMLA.
DR InterPro: IPR001227; Ac transferase.
DR InterPro: IPR004410; FabD.
DR InterPro: IPR001899; Gram_pos_anchor.

```

DR InterPro; IPR000794; ketoacyl-synt.
 DR InterPro; IPR003880; Ppanine attach.
 DR InterPro; IPR000408; Reg. chr. condens.
 DR Pfam; PF00698; Acyl. transf. 4.
 DR Pfam; PF00109; ketoacyl-synt. 4.
 DR Pfam; PF02801; ketoacyl-synt. C; 4.
 DR Pfam; PF00550; pp-binding; 4.
 DR TIGRfam; TIGR00128; fabD; 4.
 DR PROSITE; PS50075; ACP DOMAIN; 4.
 DR PROSITE; PS00606; B KETOACYL SYNTHASE; 4.
 DR PROSITE; PS00343; GRAM POS ANCHORING; UNKNOWN_1.
 DR PROSITE; PS00012; PHOSPHOTRANSFERASE; 3.
 DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
 DR Phosphotransferase; Transferase.
 KW SEQUENCE 6797 AA; 120285074050DA CRC64;
 Query Match 34.7%; Score 7221.5; DB 2; Length 6797;
 Best Local Similarity 41.8%; Pred. No. 1.5e-301;
 Matches 1784; Conservative 453; Mismatches 1373; Indels 653; Gaps 107;
 QY 52 VFRESVACEAFAFYVDMS-VEQVLRSDPAFGIDVDVQPTLFAVMIISLAAMRSQG 110
 DB 2846 VTRGAVSGADLAGAAV--KGLVRSALSEHGRFGL--VDDDDAELALVRRVLA-----S 2896
 QY 111 VEP-CAVLGSHLSEIAAAVSGSLADAAARVVTLSQAQTTL--ACTGALVSVAATPDE 167
 DB 2897 DEPOLVVRG--GEVLAAAR--LARAQSSHAVT--MPSGTVLVGTGGLGRVMAR--H 2946
 QY 168 LLEPIAWEEDNPAKLAANVNGRSTVSGARAVADL-----VA 208
 DB 2947 LV-----VEHGVNLLLVSRGP--AAGAEELVTELRHSGAEVAVEACDVTDAAYA 2997
 QY 209 DLTAQVRTMIPVDVPAHSLMYAIEERVVSGSLPTPRPSRIPIFSSVYGGRLD---- 264
 DB 2998 DLVA--RHIIISAV--VHT--AGVLDDGVESL-----TPERL--SAVLPRVDAAMN 3041
 QY 265 ---TR--ELDAAYWYNMSVTRFE-----PAARLL-- 291
 DB 3042 LHEATRLDLDAPFVSSVAGTIGSPQAVYAAGNAFLDALHHRRAAGLPAASLAWGPW 3101
 QY 292 ----OQPKTVEMSS-----PHVLTMGLOELAPDLGDTTGTADITVIMGLRGGG 338
 DB 3102 SRDGMGTGLTDVSSAPGRHARTHRGRLGFDALAAAGDAHLVPRDMASL-RAQG 3160
 QY 339 TLDDFLTLAQLRCH-----GETSATYLSARLTALSTPTQOOSLLDLVRAHTMAVLDDG 394
 DB 3161 EVPRLLGLRLTRARRRBAVGSAAAGLVGRLSGRGVERREVLDDVRAQIAVVLGHAN 3220
 QY 395 NEBTASDAGPSAPFAHLGPDVSMGVDELNRSLSKATGLRLPYTLIFDHTTPAAVAARLTA 454
 DB 3221 PRTES---TRVQDQGFSLTAVELRNRLNNAATGRLSATVAFVPTDALVDPLDE 3276
 QY 455 ALGHLDDDTAPVR--DSPSHGCTAAADDPALITGMACRFPGCVASPKDMLAASGDAT 513
 DB 3277 LFGQGEABALPAPVPSPAG-----AADDPVIVGMSCKPQGVSPDLRLVSEGDVAV 3331
 QY 514 GPRPTDGMPTQEGHADDPTQPGTFYPOGGGFLDAHPAGFGFISPREALMDPQORL 573
 DB 3332 SDPPTDGMVDESISPDPAALGTSYTRSGGFLHEAEFDPDFGMSPREALMDAQQRL 3391
 QY 574 LLETSSWEPERRAGIDPLSVSGSRTPVAGALSPDYGPRMDTASSEGADVEGHILTTGG 633
 DB 3392 LLETWEAIEERTGIDPAGLSRGTGVAGVYTDGDL-----VGDQEGYXNSGSA 3445
 QY 634 SVLSGRALYFGLEGRATVDTGCSASLVTLHLACGLRSGECTLALAGVSWSTGMR 693
 DB 3446 SLSAGRSYTFGFGPAVTVDTACSSSLVLMHMAOGLRSGECGLAAAGGTWSTPTTF 3505
 QY 694 IEFPRQGLSVDRGCKAYSAADGTGGEVGMVLVRLSDAVLIGHRYLVAVRGSVAVNO 753
 DB 3506 VEPBRQGLSADRGCKAFADAAVGVGMBGVMLVRLSDARRNGRVLAVRGSVAVNO 3565
 QY 754 DGASNGLTAPNGPAQERVIRQALNAGLSVADVVDVEGHGTGTTLGPTEAQLLATYGO 813

DB 3566 DGASNGLTAPNGPAQERVIRQALNAGLSVADVDAVEAGTGTTLDPTEAQLLATYGO 3625
 QY 814 -RAGDRPLWLGSKNSIGHTMAAGGVITKVMALREGLPTLVVDKSPQVDSAGA 872
 DB 3626 ERPEDRPLLGSKNSIGHTMAAGGVITKVMALREGLPTLVVDKSPQVDSAGA 3685
 QY 873 VRLTEAVWMP--GDAAGRRLRAGVSSFGIGTNAHVILBEPAPAGCVAGGVLEGAPGL 931
 DB 3686 VELLTSEAEWPOEGE--RAGVSSFGVSGTNAHVILBEG-----PG- 3724
 QY 932 AIVSAVSAAPVAVSAPVAESVVPVVPVVPVVSARSSEAGLRQAQALRQYAVRPOVS 991
 DB 3725 -----PDADA--APDATVTPDGCALAWLTSARNEALRQCARLSTLVAGSDALC 3772
 QY 992 LADVAGLACGRVLEHRAVVLAADEBELVQGLAAGRPDRVTTHGAPGDRGVVF 1051
 DB 3773 ARDIGSLVTGRSSPFRHRAVVMQDRDALRALASALAVGADGLLEG-ASGA--GRTAF 3829
 QY 1052 VPPGGGQWAGMGVRLASSPVFARRMQACEALAPVWMSVVDILRRDAGDAVM--ERAD 1110
 DB 3830 LFGGGSQSRILGKMGWELVATVPVADAFDAVCAALD-----EHLERLRDVWKEDEAE 3861
 QY 1111 VV-----QVLFVSVVSLAATRSYGIEBDVALVHSGSEIAAAVCGALSLKDAKTVA 1164
 DB 3882 LINGTAAQAGLFAITVALYRLAESGMRPDPVAGHSIGEVAAAHVSGVSLPDACALVA 3941
 QY 1165 LRSBALAARCGCMASVLPRAQEVQLIGERAGNLTMAAVNCPSTAVSGDAEAVDEV 1224
 DB 3942 AKRRLMOQLPSSGAMMAIRATEDEVLPRLAE--GVSLAAVNGPSVSIAGEDAVLAI 3997
 QY 1225 LAYCAGTVARARIPIVDVASHCPVOPRLRELELLGLDISPOPSGVFFSTVGATVLDTT 1284
 DB 3998 AAHFAEGKRTTLRLARSHAFHSLRMLREPARAVUTRLSFGRTPIPVSNLGLRLARE 4057
 QY 1285 TL--DAAYTRNLHQPVRFSDAVOALADGHRFVGVSVPHETVLAILEDTTEDTAEDVAT 1343
 DB 4058 QLAHADYVVRHVAEVARFADGICQALRAEGVTRLELGRPGVLSAMARASASD--DAVLA 4114
 QY 1344 GSLRGDNTRRLTLTAHTHTG-----IGPRTMNNHTNNHTNNHTNLDLP 1394
 DB 4115 PVLRRRPEBTALLGLALDLYRGANVDTVPVAGSAGRN-----ADLP 4158
 QY 1395 TYRFQHWLWLESSQAGSGSAGAGS-----GAGSGRAGTAG--GTAEVESRFW- 1443
 DB 4159 TYAFQHERFMPGSGVARPQDVNSAGLSAGHRLGLAVALBAGGGLFTGRLSVSHPL 4218
 QY 1444 -----DAVARODLETVAITLAVPSSGLD-----TVVPALSAMHRNQ----- 1480
 DB 4219 ADHVVLGSVLPQGTALVEVLRLAADBAGCDLLELTLLAARVLVPRASGAIVQVAVGEPD 4278
 QY 1481 -----HDQARITWTYQET-----WKP----- 1497
 DB 4279 EAGRRPVSVNAREBGPWTLHASGAVTSAGEVPRPDATVPRPGABRVDAUCYDVLADA 4338
 QY 1498 -LTL-PTTHQPHQWTIA-----IPETQTHPHITNLTLNHNIGTPIRLTNHTH 1547
 DB 4339 GLTYGAFGLQAMWLGSDVVAEALPES-----TDGDAVGL----- 4376
 QY 1548 TNPOLNHTLHNTROAON-----HTTGA-----ITGLSLALDETP 1585
 DB 4377 -HPALFDALHLSALGAEAGVRFSGWAGVSLHATGASHLRVRIREGAGLSVAIADTS- 4434
 QY 1586 HPHNHTPTGTLNLTLTGTHQTH-----PTLMVATTTAATTHPRDPL--THP 1634
 DB 4435 --GAPASVESLVRPLSAGQVQADRDALFKADWVPVLTDBRVRVGSGREGVPRRTYA 4492
 QY 1635 TOAQTWGLA--RTTLEHPTHTAGIDLPTRPTHTLQHTLQTLQTHQTH--OLAIRTTG 1691
 DB 4493 DIDSLEGAIVPGVIVAVRPSGAAGTVESVNAATVMALEWQAMLADDRATISLVVTRG 4552
 QY 1692 T-----HTRRLPTTLPTTHQRP----- 1709


```

QY 3796 DNAQAHLARIAGTQHEQOHTLLALVRSIATVLTGHTPTTTPDRAFRDGLFDSLTAVE 3855
DB 6615 DTAAAFAGRLTGLTSAAGREVLGAVRSQIAGVGHAEATEIDQDRAFLDGLFDSLTAIVE 6674
QY 3856 LRRRLSRTTGLRLPTLAPRHPPTTLTHLHTLOLQOPDNAAVPAVLAEIDKESALSAL 3915
DB 6675 LNRRLAAGVIGRLRPAFLPTLPALVLAHLHARIAEPVGPALGELERKEKFFGGL 6734
QY 3916 DKTDASERVTLLKSLMLRNAPQHPTAES-----ADDEKFTSATEAIFKFIINDL 3969
DB 6735 DLTEENHEQIAGRLVLAHAKMDALRDPAALAGHDSGSDPDPESASDDEVLDLDEL 6794
QY 3970 GLS 3972
DB 6795 GLS 6797

RESULT 13
ID Q93HJ4 PRELIMINARY; PRT; 4840 AA.
AC Q93HJ4;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE OLM2 protein.
GN OLM2.
OS Streptomyces avermitilis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetaceae; Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=1477403; Pubmed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinoue M., Takahashi Y., Horikawa H., Nakazawa H., Oonome T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RA "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: Deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL AB070940; BAB69193.1; -
DR InterPro: IPR001227; Ac transferase.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR004410; Fbd.
DR InterPro: IPR000794; Gram pos anchor.
DR InterPro: IPR003880; Peptide attach.
DR Pfam: PF00698; Acyl transfer.
DR Pfam: PF00106; adh_short.
DR Pfam: PF02801; ketocacyl-lyase; 3.
DR Pfam: PF00550; pp-binding; 3.
DR TIGRfam: TIGR00128; fadD.
DR PROSITE: PSS0075; ACP_DOMAIN; 3.
DR PROSITE: PSS0066; B_KETOACYL_SYNTHASE; UNKNOWN_3.
DR PROSITE: PSS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE: PSS00012; PHOSPHOANETHEINE; UNKNOWN_2.
KW Phosphoanethine.
SQ SEQUENCE 4840 AA; 504325 MW; 169BD3589AD284F CRC64;

Query Match 34.6%; Score 7217; DB 2; Length 4840;
Best Local Similarity 39.3%; Pred. No. 1.5e-301;
Matches 1799; Conservative 475; Mismatches 1419; Indels 878; Gaps 97;
QY 6 GGGEPRAAGEVLCVADADGCVFVPPGGPQMPCKGRLLDASDVFRSVCACAPAA 65
DB 542 GBDPFRVGRRTV-----PSGPVAFLEPTGGSQRAAGRGRLYTFPVFAFAPEVDRLD 596
QY 66 PYVDMSVEOVLROSDPAAGLDVVDVVOPT-----LFAVMIISLAALRSQGVPCAIVGH 119
DB 597 POLAHPIDIL---DRLVVDGLPLLEQIGYALAAVFALEVALFRLEHAAGVVPDVLGH 652
QY 120 SLGEIAAAHVSGSLADAAARVVTLSQAQTTLAGTALVSVATPDDELPRIAPTEDN 179

```

```

DB 653 STGEIAAAACAVLSIDDACITLVAARGFVQTLRTDGAANVAVETSETEIAELIA--RAVY 710
QY 180 PARLVAANVGRSTVSGAREAVADLVADLTAAQVRTMIPDVPAHS-----PLMYALE 235
DB 711 GDAITTAIVNGPRSVVSGDADVERVVRARCEBGRKTKLPLSHAHSVHNEEMLARLR 770
QY 236 ERVVSGLPTTPRPSRIIPHSVGTGRLDTRLEIDAAVWYRNNSVTVRFPAPARLLIQG- 294
DB 771 E-VAEGL---EHRPPIPMVSNLTGA.PVPEANWPEYVWROYRETVRFHDGIRMLDAQV 826
QY 295 -----PKTFV-----EMSPHVL----- 307
DB 827 TACLELPGCALIAMARESLGDPGREVALPTLQHDRPEPATFLSALAEIHHVAVAVMD 886
QY 308 -----TWGLQE----- 313
DB 887 TPLAARGATSTDLPTAFAFORERYWLPATPAAPRTAEPADRTRGRVRYGVGWTYPVDVDTA 946
QY 314 -----LAPDLGDTTG-----TA-----DTVIMGT 332
DB 947 RPTGTWLVVAPD-GGSDGDGHGDGARVAALGEBRAVLTAGGGLAPALRLPBDRAVAVY 1005
QY 333 LR-----RQGT----- 339
DB 1006 LALPAESATRWTRLAPRSALRGARPASVRHQAPVSVTEAVPGCAPALMGLGRVPHS 1065
QY 340 ----- 339
DB 1066 GGGVVDLPAGLDARSRLRLAVALAAPGODELAVRPGVFAKLLPAPRPRQRANRTG 1125
QY 340 ----- 339
DB 1126 TVLTGDIAGVTPPELLQVYSGGARTLLAHRPGTPLPDAVLAAGCCTYEMDLGLP 1185
QY 340 -----LDHF----- 343
DB 1186 ETASVAVVHLSGIDPDGSPDPAITAGAPAVODRLRTVRLSELPHDADLDIAFVLS 1245
QY 344 -----LTSIAQLR-----GHGTSATL----- 360
DB 1246 TVAGTWGAEDLVHTVAHAALSLAERRTLDRPGTCVGWGPEVTPADTGASBPGLPLP 1305
QY 361 -----VLSARLTALS-----PTQOQ 375
DB 1306 PADLAVSALAEALDADDPVCAVVDVMSLFHVLTGRGRPLVADLPEVRALPAPRTAVY 1365
QY 376 SL-----LIDLVRATMAVLNDGNERTASDAG--PSASPAHLGFSYWGVELRR 424
DB 1366 VLAPADREHOLELTVRAQAAVLAGH-----ASGAALDPVAPRFDLGFESLAAVAFRDR 1419
QY 425 LSKATGLRLPVTLIPHTTPRAVAAARLRTAALG---HLDEBPAPVDSGSGCTAAAD 481
DB 1420 VAAETGLRLPATLVFPHTPAQAVALLAELSGVRHBE-VAP-----TADHDD 1468
QY 482 PLAIIGACRPFQGVSPDLMELASGGDAIGPPTDGM--PTBORAODPTQGTGF 539
DB 1469 PLAIIVMACRYRFGVGRPEDLMDLVHCRDVGDPFTDGMDLPALRLARP-----L 1521
QY 540 PGGGFLDAAHFADAGFGISPREALAMDPOQLLLETSMFAERAGIDPLSVRGSRTGV 599
DB 1522 AMKGGFLPADAFDAFPGISPREATAMDPOQLLLETSMFAERAGIDPLSVRGSRTGV 1581
QY 600 FAGALSFDYGRPMDDTASSGADVEHILITGTTGVSLSRIAYRFLBEPALITVDTCGA 659
DB 1582 FAGAASDVGVAL-----EGAEADGVLTTGTAASVISGRIRIYTTGLHBPALVVDVPAAS 1636
QY 660 SLVTLHLACOSLRSGECTALAGVSVMSLTGMFIEFSRQGLSVDRCKAYSAADGTG 719
DB 1637 SLVALHMAVGALHOGECDLATGVSISTPAPFVDFAPQSGLAADGRCKAPASADGTR 1696
QY 720 WGEVGMILLVERISDAVLRVLAHVRSAAVNODASGLTAPNGPAQERYIROLAANA 779

```

Db 1697 MAEGVGLVVERLSDBARNGRVLATVRSAGVNSDAGNSGLSAPNGAQRVIRQALAAA 1756
 Qy 780 GLSADVDVVEGHTGTTLDPRLEAQAALLTYGQ-RAGDRPLMTGSLKSNIGHTMAAGV 838
 Db 1757 GLSADVDVAEAGHTGLTDPRLEAQAALLTYGQDRAGOPLMGSLKSNIGHTSAGAAV 1816
 Qy 839 GGVIKMMALREGVLPRTLVNDKPSPOVDMSAGAVRLITEAVPPRGDAAGLRAGVSS 898
 Db 1817 GGVIKMMALREGVLPRTLVNDKPSPOVDMSAGAVRLITEAVPPRGDAAGLRAGVSS 898
 Qy 899 GIGSTNAHVILEAPRAAGGVAGGVLEGARGLAISVAESVAPVAVSAPVAVSV 958
 Db 1875 GIGSTNAHVILEAPRAAGGVAGGVLEGARGLAISVAESVAPVAVSAPVAVSV 958
 Qy 959 VAVVAVSAPVAVSAPVAVSAPVAVSAPVAVSAPVAVSAPVAVSAPVAVSAPVAVS 1018
 Db 1917 VAVVAVSAPVAVSAPVAVSAPVAVSAPVAVSAPVAVSAPVAVSAPVAVSAPVAVS 1018
 Qy 1019 ELVQGLGALAAEEDRRVTTGHAEPGDRGVVFPVPGQGMAGVRLIASFPVARRM 1078
 Db 1974 ELVQGLGALAAEEDRRVTTGHAEPGDRGVVFPVPGQGMAGVRLIASFPVARRM 1078
 Qy 1079 QACEBALAPVWDSVVDILRRDAGDAVWRADVVQVLFVWSVLAALMSYGIERDAVL 1138
 Db 2034 QACEBALAPVWDSVVDILRRDAGDAVWRADVVQVLFVWSVLAALMSYGIERDAVL 1138
 Qy 1139 GHSGEELAAHVCCALSLKDAKTVARSLAAVRRGMAVSUPLPAQVEOLIGRMA 1198
 Db 2091 GHSGEELAAHVCCALSLKDAKTVARSLAAVRRGMAVSUPLPAQVEOLIGRMA 1198
 Qy 1199 GRLVAAVNGPRSTAVSGDAEAVDEVLAYCAGTVARRIPVDYASHCPVQPLREELLE 1258
 Db 2150 GRLVAAVNGPRSTAVSGDAEAVDEVLAYCAGTVARRIPVDYASHCPVQPLREELLE 1258
 Qy 1259 LAGDISPQSGVPESTVEGWTDTTLDAAYTRNLHQRVSDAVOALADGHRVVE 1318
 Db 2210 LAGDISPQSGVPESTVEGWTDTTLDAAYTRNLHQRVSDAVOALADGHRVVE 1318
 Qy 1319 VSPHPTVPAIEDTEDETAADVTAIGSLRGDNDTRFLALANTHTTGTGPTTMHNY 1378
 Db 2270 VSPHPTVPAIEDTEDETAADVTAIGSLRGDNDTRFLALANTHTTGTGPTTMHNY 1378
 Qy 1379 THNHTHPRHTHLDLPYFPOHNYWLESSOPGAGSGSAGSAGSAGSAGSAGSAGS 1438
 Db 2328 THNHTHPRHTHLDLPYFPOHNYWLESSOPGAGSGSAGSAGSAGSAGSAGSAGS 1438
 Qy 1439 ESRFWDVAADOLEFVATTAAPRSAGLDVVRLASMRHNDQADINTWTYQETWKL 1498
 Db 2361 ESRFWDVAADOLEFVATTAAPRSAGLDVVRLASMRHNDQADINTWTYQETWKL 1498
 Qy 1499 TLPTTHQPHOTWILAIPEOTHHNHILNHLNHHGITRPLTLNHTHNPQHHLHTL 1558
 Db 2421 TLPTTHQPHOTWILAIPEOTHHNHILNHLNHHGITRPLTLNHTHNPQHHLHTL 1558
 Qy 1559 HTRQQAONHTTGAIT-----GLSLLA 1580
 Db 2461 HTRQQAONHTTGAIT-----GLSLLA 1580
 Qy 1581 LDETPEHNPHTTGLTLNLTQTQHTQHPRLMYATTNATTNBNDRPLHPTQACT 1640
 Db 2521 LDETPEHNPHTTGLTLNLTQTQHTQHPRLMYATTNATTNBNDRPLHPTQACT 1640
 Qy 1641 GLARTTLLEHTHAGIIDLPTTPHTLOHLOTLQORHNOQALRTGHTHRTPLT 1700
 Db 2581 GLARTTLLEHTHAGIIDLPTTPHTLOHLOTLQORHNOQALRTGHTHRTPLT 1700
 Qy 1701 TLPTTHQPHOTWILAIPEOTHHNHILNHLNHHGITRPLTLNHTHNPQHHLHTL 1754
 Db 2641 TLPTTHQPHOTWILAIPEOTHHNHILNHLNHHGITRPLTLNHTHNPQHHLHTL 1754
 Qy 1755 HAQHLTQLOQKIHLLTTCDSNPDOLOQLNTIPPO-HPLTTVHTHAGIILDAATLTN 1813
 Db 2693 HAQHLTQLOQKIHLLTTCDSNPDOLOQLNTIPPO-HPLTTVHTHAGIILDAATLTN 1813
 Qy 2693 GADALVTELTLGLVKVLLTSCDLADPAQVAMVAEORDDBPGAVHTHAGGGLGLAE 2752

Qy 1814 LPTQOLNNTVRAKASHHLLHOUTHTPLTAFLVSSAATFGAPQOANYAANYLIDAL 1873
 Db 2753 LPTQOLNNTVRAKASHHLLHOUTHTPLTAFLVSSAATFGAPQOANYAANYLIDAL 1873
 Qy 1874 AHRHTHNPHTTGLTLNLTQTQHTQHPRLMYATTNATTNBNDRPLHPTQACT 1933
 Db 2813 AHRHTHNPHTTGLTLNLTQTQHTQHPRLMYATTNATTNBNDRPLHPTQACT 1933
 Qy 1934 PYVVIADIDMSK-----IHTSQTSDVSAAREPRAVORPT---PAEHTLAHQS 1984
 Db 2872 PYVVIADIDMSK-----IHTSQTSDVSAAREPRAVORPT---PAEHTLAHQS 1984
 Qy 1985 ADOQAALLETVRHVAAYLPHADPKAIAPDOSRALGFSLTAFERHLLIKATGLRLV 2044
 Db 2930 ADOQAALLETVRHVAAYLPHADPKAIAPDOSRALGFSLTAFERHLLIKATGLRLV 2044
 Qy 2045 SLVDFHPTPAKLAHVLONOLRGTAESA-PSAAVTAASVTEPIAIVGACRFPQGVTS 2103
 Db 2990 SLVDFHPTPAKLAHVLONOLRGTAESA-PSAAVTAASVTEPIAIVGACRFPQGVTS 2103
 Qy 2104 ADFPWLISSEODALIGFPTDRGMDLTLD-PDRPHPTCTYRNGGFLYDAGHFAEP 2162
 Db 3047 ADFPWLISSEODALIGFPTDRGMDLTLD-PDRPHPTCTYRNGGFLYDAGHFAEP 2162
 Qy 2163 GISPREALAMPQORLLLETAETIEHAGINPTLHGTPTGVTGNTGODYALRVHNAQ 2222
 Db 3107 GISPREALAMPQORLLLETAETIEHAGINPTLHGTPTGVTGNTGODYALRVHNAQ 2222
 Qy 2223 STDGALVTGAGVISIGRISYTFEGEPVAVSDTACSSSLVALHILACQALRAGECMA 2282
 Db 3167 STDGALVTGAGVISIGRISYTFEGEPVAVSDTACSSSLVALHILACQALRAGECMA 2282
 Qy 2283 GGVTVWSSPCAFEPERORGLAADGHCAPSAADGTGNGEGVGMILVERLSAHNNGR 2342
 Db 3226 GGVTVWSSPCAFEPERORGLAADGHCAPSAADGTGNGEGVGMILVERLSAHNNGR 2342
 Qy 2343 VLAIVGSAVNOGANGSLTAPNGPQOQVIRIOLANAGSADVDVAEAGHTGLTLD 2402
 Db 3286 VLAIVGSAVNOGANGSLTAPNGPQOQVIRIOLANAGSADVDVAEAGHTGLTLD 2402
 Qy 2403 IEAQAALLATYQODRAGEBPLMGLSVKSNVGHOTOAAAGVAKMVALRHGLLPTL 2462
 Db 3346 IEAQAALLATYQODRAGEBPLMGLSVKSNVGHOTOAAAGVAKMVALRHGLLPTL 2462
 Qy 2463 EBPHTVMSAGAVOLLTEVPWPGEGRLRAGVSSFGVSGTAAHITLLEAPRDVPG 2522
 Db 3406 EBPHTVMSAGAVOLLTEVPWPGEGRLRAGVSSFGVSGTAAHITLLEAPRDVPG 2522
 Qy 2523 --PAGE-----DAGSDDEAAGSPGV-----WPLVSAKSOPLRAQAOLH 2563
 Db 3464 --PAGE-----DAGSDDEAAGSPGV-----WPLVSAKSOPLRAQAOLH 2563
 Qy 2564 AMLTDPHGLDLADVGTYLAAHARAVPDRATLAAADRTFLQALQALAGEPHAVIHSSA 2623
 Db 3524 AMLTDPHGLDLADVGTYLAAHARAVPDRATLAAADRTFLQALQALAGEPHAVIHSSA 2623
 Qy 2624 PGGTGGEAAGKTAFCGSGOGRPMAAGLHTHVPRAALNDICTHNDPHLDPL 2683
 Db 3571 PGGTGGEAAGKTAFCGSGOGRPMAAGLHTHVPRAALNDICTHNDPHLDPL 2683
 Qy 2684 LQONDNDNEADALLOOTVYAORALPAFOVALHRLITDGVHTTPHYAAGSLGEITAAH 2743
 Db 3630 LQONDNDNEADALLOOTVYAORALPAFOVALHRLITDGVHTTPHYAAGSLGEITAAH 2743
 Qy 2744 AGILLTDAATTLTORATLMQTPR-GTMTTLTTPHTHNTLHNTAHENDLAAINTPTS 2802
 Db 3687 AGILLTDAATTLTORATLMQTPR-GTMTTLTTPHTHNTLHNTAHENDLAAINTPTS 2802
 Qy 2803 LVISGTPHTVQHTTTCQOQKIKTKTLPNNHAFHSHHTPILNOLHONOTLTYHPPH 2862
 Db 3747 LVISGTPHTVQHTTTCQOQKIKTKTLPNNHAFHSHHTPILNOLHONOTLTYHPPH 2862


```

QY 309 ----- 308
Db 899 AEWAGVFECTAREVGCCEVLEPTVAFERERFWLVBEGSAGSGVSGMMGFLMEAVE 958
QY 309 ----- 308
Db 959 CGDAGVVASLGVDEGASLGAVVSALGEMGRVHRHEREVVDGMYREYWRPVSGGVGLS 1018
QY 309 ----- 308
Db 1019 GAMLVSEGEAPVDVVAEGIERCGARVVRVEAGCVSREVLAGHLREAVDGEAVGVV 1078
QY 309 ----- 308
Db 1079 SLVMSGGVVQAGVASVGLVQALDGVGARLWCTTGAVSVGGRDAVWGPASGVWGLG 1138
QY 309 ----- 308
Db 1139 RVWGAEPDRWGLVDVPELVDERVVDGLVGLAGVGGGSEBPVSSGAFFVRLVRA 1198
QY 309 ----- 317
Db 1199 LEEVAEREMRPRGTVLVGTGELGAHVANMARMAHGAHLLVSRKESAGVEELRAD 1258
QY 318 L-----GDTTGTADTVIMG 331
Db 1259 LMGIGARVSVACDADREALAEVLSAVPAECPLGVVVAAGVVDGVLGLSERVYT 1318
QY 332 TL-----RKGCG 338
Db 1319 VLRAKALAAMNHELTRGADLSCFVVSAAATFPGAGQGSYAAANVVEALVHRREG 1378
QY 339 ----- 361
Db 1379 LPGLAVAMGPWAGGMAEGAVGQRRRGLAMTETALVALGQALDHDETCTVADLDWD 1438
QY 362 -----LSARLTLSPTOQSLDL 381
Db 1439 RFTANSLPGSRSLPSLIDIPEARLARETTGLDTATASPDSPSALKMDTAEGRALLDL 1498
QY 382 VRANTMVLNDGGERLNASDGPASPAHLGFDVSMVELENNLSKATGLRLPTLLFDH 441
Db 1499 VRITAAVTL-----GHSTPTAVRPRAPRDLGFSVSAVELNRRNATVGLLLPTLLFDY 1554
QY 442 TTPAVALARL-----TALGHLDEPTAPVPSGHCCTAADDIATIGMACRPPGVS 498
Db 1555 PTPBALAGYLKEQLEBAGGQORDIAPVPAS-----RVVDEIATVGAHCRPPGES 1608
QY 499 PKDLMELASGDAIGPPTDRGWPTERQNAQDPTQPTPYQCGFLHDMAHFDAGFG 558
Db 1609 AEDLMELVASGRDAVGEPPVDRGWDVEAFYDPEBGRAGSSYTRRGFLBGAEPDAGFG 1668
QY 559 ISPREALAMPQOQLLETSWEAFERAGIDPLSVRSRTGVFAGALSFDYCPMDTASB 618
Db 1669 ISPREALAMPQOQLLETSWEAFERAGIDPATLRGTTGVFACMCQDVAULVRRATE- 1727
QY 619 GAADVEGHILNLTGTVSLSGRIASFGLEBPAITVNDGCSASIVTLHLAOCULSBSGCTL 678
Db 1728 ---DLEGANTGLSSSVTSGRVAITLLEBPAVITVDACSSIALHLAOCULSBSGCTL 1784
QY 679 ALAGVSVMTLGMFIEFSRQGLSVDRCKAVSAADGTGMBGVMLVLERLSADVRL 738
Db 1785 ALAGVTVMTSPGAFVEFSRQGLSPDRCKAVSGADGVGMAGVGLVLERISEAERR 1844
QY 739 GHRILAVYRGSANVODGASNGELTAPNGPAQERVIRQALANAGLSVADVVEGHGTTL 798
Db 1845 GHRILAVYRGSANVODGASNGELTAPNGSQORVIRQALACAGLSVADVVEGHGTTL 1904
QY 799 GDPTEAOLATATYGO-RAGDRPLWGLSKNIGHTMAAGVGYIKWMMALREBVLRTL 857
Db 1905 GDPTEAOLATATYGOQSSGERPWLGSYKSNIGHAQAAGVAYIKWMMALRAVLRRTL 1964
QY 858 HVDKPSQDVMSAGAVRLTTEAVPMPGDAGRLRRAGVSPFGIGTNAHVILBEAPAGG 917

```

```

Db 1965 HVDKPSQDVMSAGVAVLDADEVWEPG-VEGRILRRAGVSAFVSGTNAHVILBE--ASCG 2021
QY 918 CVAGGVL-BGAGLMAISVAESVAAPVAVSAPVAVSPVVPVPSVARSAGEAQRQ 976
Db 2022 ADGAGRLQELGPGV-----VSGSEVWVVSARSELALRGQ 2058
QY 977 AEALRQYVAV-----PPDVSLADVAGLACGRAVLLEHRAVVLADREELVQGLALAGE 1031
Db 2059 ARRLRGVAVVGGADGCVSPAGVGRALVSESVFHHRAVVADEDEFLHALDALAGA 2118
QY 1032 PDRVTTGHAPEGDRGVGFVFPQGGQMAQWGVRLIASSPVARRVQACEALAPWVD 1091
Db 2119 PTAGVVOGVGAPADGKIAMLFGQGGTHMEGAQELLGSSPVPAAQMSDCAQALEPYLDW 2178
QY 1092 SVVDILLRDAGDAVWERADVVOVULFSVWVSLAALRSVIEEDAVLGSQGEIAAHV 1151
Db 2179 SLLDVLRGADPADPDLQVDVVOVULFAVWVSLAALRSVGVHDDAAGHSQGEIAAVVA 2238
QY 1152 GALSIXDAKTVALLRSRALLAABVGRGMAVPLPAQEVQELIGERVAGRLMAVANGPRS 1211
Db 2239 GALSIXDAKTVALLRSRALLAABVGRGMAVPLPAQEVQELIGERVAGRLMAVANGPRS 2297
QY 1212 TAVSGDAEAVDEVLAACAGTVARARI PVDYASHCHVORLREBELLELGDISPQSGVP 1271
Db 2298 AVVSGEPAVDALVEELSHEDVPARLMDWMASHSPQVEAIOGRLELILAPARTGDVP 2357
QY 1272 FFSVTEGTMVDTTLTAAVWYRRLHQVPRPSDAVQALADGHHVFEVSPHPLVALIED 1331
Db 2358 FYSVTVGERIDGELLDADVWYRRLHQVPRPSDAVQALADGHHVFEVSPHPLVALIED 2417
QY 1332 TTEDTAE-DVTALGSLRGDNDTRREFLTALANTHTTIGITPTTHNHTYHNHTPHPT 1390
Db 2418 TIDEMGDLDBLVVSGSLRGSGGLARFLMSVABLPVSGVA--VMSGVFGSVGRVAGCG 2475
QY 1391 LDIPTYPQOHYWLSSOPGAGSGSAGAGSGAGSGRAGTAGTAEVBSRFMDAVAROD 1450
Db 2476 VELPTVAFERERFWLVE-----GABRGSGVSGVSGOMG-- 2507
QY 1451 LETVATTLAVPRAGDVTVPBALSAHNRHQHDAINTWYQOCTWKPRLPTTHQPHQW 1510
Db 2508 ----- 2507
QY 1511 LIAIPETQHNPHITNLITLHNHGTPLRLTLNHTNPNQHLNHTTROQAOHNHTG 1570
Db 2508 ----- 2507
QY 1571 AITGLSLALDETPRPHNHPRTGLNLTLQTNHTQHPRLPVATTNATTNHPNDP 1630
Db 2508 ----- 2507
QY 1631 LTHPTQATWGLARTLLENHTHTAGIIDLPTPTHTNHLQTLQTLQPHNQOLAIRTT 1690
Db 2508 ----- 2507
QY 1691 GHTRLRLPTTLPTHQRPRTPHGTTLLTGSGALATLHNLHTHQPHLLTSRTG 1750
Db 2508 ----- 2507
QY 1751 PHTPHOHLTTLQOKGHLTTTCDSNPDLOQLANTIPRONPLTVTINHAGIIDAT 1810
Db 2508 -----RGSMRLDC- 2524
QY 1811 LTNLTPTQNLNVLRARAHSHNLHLQTLQHTPLTAPVLVSSAAATPGAPQOANYAANAYL 1870
Db 2525 ----- 2524
QY 1871 DALAHHRHTNHLPATSIAMGTWQNGLADSKARAVILDRGRFMSPELATAVTOAIAD 1930
Db 2525 ----- 2524
QY 1931 TERPVVVIADIDMSKIHTSQTSDLVSAAREPRAVQPTPPAELHKTLAHQTSADORA 1990

```

Db 2525 -----LAGLD-----PAAQAE----- 2536
 Qy 1991 LLELVDRHVAVLRHADPKAIAPODFRALGFDSLTAVERFNLILKATGLRLPVSLVFDH 2050
 Db 2537 VLDELVLTHAAAVLCHGTADAVPERAFRDLGFDLTVELRLNRLNTATGLRFPRTLVFDH 2596
 Qy 2051 PTPAKLAVHLONOLRG-----TAESAPSAAVTAASVTERPIALVGNACRPPGVTSAD 2106
 Db 2597 PRPAVLAHIIHQSLSGSPFTTGTLALALRAPRVVDPERIALVGNACRPPGVTSAD 2656
 Qy 2107 FMDLISEBODAGFPPDRCHMDLDTLYDRPDHGTCTYRNGGLYAGHFDAPFISIP 2166
 Db 2657 FWEIAGGRDAVGFPPDRGVDVEAFYDPEGRAGTSTRCGGLQGAADFAGFISIP 2716
 Qy 2167 REALAMPORLLLETAMETIEHAGINPHTLHGPTGTCNGDVALRVHNAQSGTDS 2226
 Db 2717 REALAMPORLLLETAMETIEHAGINPHTLHGPTGTCNGDVALRVHNAQSGTDS 2776
 Qy 2227 FALTGTAGSVISGRISYTFGEFEGBAVSDTACSSSLVALHLACQALRAGECSMALAGVT 2286
 Db 2777 YALTGVSNSVSGRLSYTFGEFEGBAVSDTACSSSLVALHLACQALRAGECSMALAGVT 2836
 Qy 2287 VMSPPGFVFEFSRGRGLAADGCHCAFSAAADGTGMEGCVMLVERLSDAHRNGHRLAV 2346
 Db 2837 VMSPPGFVFEFSRGRGLAADGCHCAFSAAADGTGMEGCVMLVERLSDAHRNGHRLAV 2896
 Qy 2347 VRGSAVNOQDASNGLTAPNGPSOORVIROLANAGLSAGDVDAEHAHTGTLLDPIEAQ 2406
 Db 2897 VRGSAVNOQDASNGLTAPNGPSOORVIROLANAGLSAGDVDAEHAHTGTLLDPIEAQ 2956
 Qy 2407 ALLATYQODRABEGPMLGVSXSVNGHTOAAAGVAVIKVMALRHGLPRTLHVDESP 2466
 Db 2957 ALLATYQODRABEGPMLGVSXSVNGHTOAAAGVAVIKVMALRHGLPRTLHVDESP 3016
 Qy 2467 HVMDSAAVQULITETVWPGEGRLRAGVSVFSGSTNAHILEAR-----ADVPFG 2520
 Db 3017 HVMDSAAVQULITETVWPGEGRLRAGVSVFSGSTNAHILEAR-----ADVPFG 3075
 Qy 2521 -----GP-----PAGEGAGSD-- 2532
 Db 3076 EAHYIVQARVEARHNVDAATGPRADLTERRQEAARECVADAVTMSAEARECVADAMSEMS 3135
 Qy 2533 -----DEMAAGSPGVW--PMLVSAKSOPALRAQALNHLTDHPGL 2572
 Db 3136 ABCVAEAVSDKSAARECVADAMSDKRALRLPMLLSAKSBRALRGQARLRQFARASDA 3195
 Qy 2573 DLADVGTTLAHARAVERPHRATLLAADDTGLOALQALAGEPRHVAHSG-ARCGSTGE 2631
 Db 3196 RPADVAAHAAQRSVPHRAVAVVEDRDFLOALDALAEGRSADGLIEGSGVPRGSHGR 3255
 Qy 2632 AAGTATIGSGGOGORPGMAHGLYHTRPVRAALNDICTHLDHRLPLRLTONN-- 2689
 Db 3256 RRGSTALHFAAGQTORVGMGRQLYAHRPAADADOLVABLDGLDRLRLHNASADLA 3315
 Qy 2690 DNEBAAALLOOTRYAORALFAQVALHRLTDGYNITPHYUAGHSLSEITAMHAGILTL 2749
 Db 3316 DVADAADVLRTRTAQORALFAVQVALFRHL-ERLGVADAFVAGHISIEBLAAHYAVGLPL 3374
 Qy 2750 TDAATTLTORATLMQTRP-STMTTLHTPHNITLHNLTAENDLALAIANTPISLVISGT 2808
 Db 3375 AAMCRLVAAGRIMEOLAFCGAMVAVRASEABARQALDREARVSVAUVGPRASVPSGA 3434
 Qy 2809 PHTYONHTTLCOOGIGTKTLLPTNHAHSPHTNPIILNOLNONTGLLYUNPHRLTANT 2868
 Db 3435 EDEVGNAADPFAEGRKVKRLRTGHAHSPRLMDPMLSEFOOVAASLTYSERATWVSTLT 3494
 Qy 2869 -----PDDOLTPHYWTOOARNTVDYATTTOTLHONGVYVIELGRDNTLTTLTHNN-- 2921
 Db 3495 GDVIAAGELSDPEYVWQVARTVAFGDAISRLHTDGVRTFMEIGRDPDGLSLABESECLEAT 3554
 Qy 2922 -----PNEPTT-----TLTTPRHNNHQTLLNLAKTTTTPHNNHTHNDONPHN-- 2968
 Db 3555 ADSHRADDGTGTPOENLILPLLRPSRPPGTLTGLARL-----HTHGA 3598

Qy 2969 -----THLDLPTYRFOHNNHWTSESTOPGAGNVSAAGIDPTEHRLGATLE 3013
 Db 3599 AAVNWPALPERBARHLDLPYAFDHHRYWDTSAHGPDLSAAGLGTAGHLLSASVA 3658
 Qy 3014 LATDGALLAGRLSLRSHPLADHAGVGVLLSGATFSLALHAGTYVCGDRVDELTLHA 3073
 Db 3659 LAESOBLLFTGRSLSTHPLADHAFGVTVLLPGTALILELAVAAGDEVCCGYEBELTLRT 3718
 Qy 3074 PLVVPVDSGVSVGVAAADG-----EGRRLVSVAR--GSGSAGGSGAGSVWTCNAS 3125
 Db 3719 PLVVPVDSGVSVGVAAADG-----EGRRLVSVAR--GSGSAGGSGAGSVWTCNAS 3778
 Qy 3126 GVLVEAAGGVVVDGAGVPPRGAVALVDGVRDLAAGCVLGFVSGSLRAVMEBGD 3185
 Db 3779 GVLVTGAR--PHEHNOEPMPRADAPVLDGVEBGLAAGLGCGVFPGLRVRMGRDE 3836
 Qy 3186 LLAECVCLPEBAMG--DAAGFGLHPLALDGVOPSLVLLPGTGFEAGFGEGRVPAV 3242
 Db 3837 VFAVVTLPESBEGQAADARVALHPLALDAAHPV--VLHNEGDAADGHW-----LPFS 3890
 Qy 3243 MCGVSLHRAVGTVGRVSAVGRGGRGAASVVVGDEGVPVSVDRLELRPVDMQOLA 3302
 Db 3891 MCGVSLHRAVGTVGRVSAVGRGGRGAASVVVGDEGVPVSVDRLELRPVDMQOLA 3947
 Qy 3303 VSVSAGRGSGLYAVQ-----MAEVCB-----VPCGQAMAMHEDVGE 3339
 Db 3948 --ARTGVHDLPRIEHRPLHPTTPARTADMALIGARRTAVLERNGASWQAYDPA 4005
 Qy 3340 -----SGGFPVGVVLR-C-PD-----AGAGGGGGGGGGGGEV 3373
 Db 4006 LAELALGAAPAPGMVVISCEPDDASAPTDALSATDSAPASGAPADSTALADATRQA 4065
 Qy 3374 VGVVLVGVCGMLERFAGSRVLYVVRGAVVVGPEBGPDVGVASVWGLVRSQAQBPR 3433
 Db 4066 TTRVLLLOEVADEBLACRLALLTHGVTATRPDEPSDLAAHVAWGLVRSQVTEPR 4125
 Qy 3434 FVLLDLDTDTGTDLTGAGAGMVGDCGRVAVAACEPOLAVGERLLAARLXLESSE 3493
 Db 4126 FVLLDLDTDTGTDLTGAGAGMVGDCGRVAVAACEPOLAVGERLLAARLXLESSE 4170
 Qy 3494 VPAORSQDTRARRSDVPAORSGVPAARSVDVSGREVLVWLSGGSVLYVTGTCVGAVA 3553
 Db 4171 QPSTTAD-----DASRNALPLLAEBEPOALRNGAVRIPRMTRVPROR 4199
 Qy 3554 RHLAGVGRDILLVSRGRDAPGARGLAELAALGAEVRIYACDVGERRVRLLEGV 3613
 Db 4200 RHLATAGVGRDILLVSRGRDAPGARGLAELAALGAEVRIYACDVGERRVRLLEGV 4259
 Qy 3614 AGCPITGVVAAAGVLDATIASLTPRLGTVFAAKVDAALLDELTRGMELSAFVPSA 3673
 Db 4260 ACPRLKAVNHTAGVDDGILTSITPRMEVULHAKFAGAHNHLTRDAGLTFTYVPSA 4319
 Qy 3734 RSGLHMSVTPDALALDADALADR-----PVLPRDL-----RPARPLRPLDOL 3779
 Db 4380 RSGLHMSVTPDALALDADALADR-----PVLPRDL-----RPARPLRPLDOL 4436
 Qy 3780 PATRRRTTRTTTGADNG-----AQLNARLAGOTHEOHT--TLALVRSHTATVLCNT 3832
 Db 4437 RITARRS-----ADHGATATPTATLBERLHAGLITIPRAOTGLLELVYTHAAAVLGNR 4488
 Qy 3833 T-----DGTIPRDAFRDLGFDLTAVERLNRRLSRTTGLRLPTTLADPHNPT 3881
 Db 4489 TSAVTAADGALPDLDVAPADEFRDLGFDLTAVERLNRRLSRTTGLRLPTTLADPHNPT 4548
 Qy 3882 LTHNHTOL-----QORPDVA 3899
 Db 4549 LADHLATRLTAEAGTDEPARAAAA 4574

RESULT 15

OQ3NM6

PRELIMINARY; PRT: 10917 AA.

AC OQ3NM6; 01-DEC-2001 (TREMBLREL. 19, Created)
 DT 01-DEC-2001 (TREMBLREL. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLREL. 21, Last annotation update)
 DE AmphC.
 GN AMPHC.
 OS Streptomyces nodosus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyocineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=40318;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Caffrey P., Lynch S.V., Flood E.M., Finnian S.M., O'Leary M.,
 RT "The amphotericin biosynthetic gene cluster from Streptomyces
 RT nodosus."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP357202; AAK73514.1; -
 DR InterPro: IPR001227; Ac_transferase.
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR InterPro: IPR004410; FabD.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR InterPro: IPR003880; Peptide attach.
 DR Pfam: PF00698; Acyl_transf. 6.
 DR Pfam: PF00107; adh_zinc. 1.
 DR Pfam: PF00109; ketoacyl-synt. 6.
 DR Pfam: PF02801; ketoacyl-synt. C; 6.
 DR Pfam: PF00550; PP-binding. 6.
 DR TIGRFAMs: TIGR00128; fabD. 6.
 DR PROSITE: PS00075; ACP_DOMAIN. 6.
 DR PROSITE: PS00059; ADH_ZINC. UNKNOWN 1.
 DR PROSITE: PS00060; B_KETOACYL_SYNTHASE. UNKNOWN 6.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN 5.
 KM Phosphopantetheine.
 SQ SEQUENCE 10917 AA; 1132905 MW; 15AC5956B5610A1 CRC64;

Query Match 34.4%; Score 7168; DB 2; Length 10917;
 Best Local Similarity 41.5%; Pred. NO. 5.9e-299;
 Matches 1742; Conservative 475; Mismatches 1374; Indels 603; Gaps 93;

OY 154 GTGALVSAATPDELPRIAWTEDNPARLAAVANGPSTVV-----SGAEAAVADIVAD 209
 DB 6552 GTGELGAVLAR-----HLVAEHGVRELVLVSRGGAAGAAELVAE 6992
 OY 210 LTTAAQVTRMIPVDVPAHSPMLVAIEERVSGL-----LPITPSPRIPEHSSVT 259
 DB 6993 LAESGATATVACVDTBRAAAELVAHPVSAVHSGVLDGLIASLTPERL---SGVL 7049
 OY 260 GGRILD-----TRELD-----AAVYRN--MSSTVFEPPA-- 287
 DB 7050 RPKVDAAWNLEHARRDLDAFVVFSSVAGVFGNAGQANVAAAGAFDALLHEHRAAGLP 7109
 OY 288 RLILQOGP-----KTFVEMSPHVLTMLGLOELADLG-----DTGTADIVINGTL 333
 DB 7110 GLSLAMGWEOGTGCTGTSIDAELERLASSGVPLAVEGGLALFDALSTEDAAVLPV 7169
 OY 334 R-----RGQGLDHLFSLSLAQLRGH-----GETSATTVLSARLTALSPTOOGLLDLV 382
 DB 7170 RLDDPALRGGEVPEPLRSLLRARSRAAAGSATATG--LDRDLAVDAVRRREVLDV 7228
 OY 383 RAHTMAVINDGNERTASDAGPSASFAHLGFDVSWGVELRNRLSKATGLRLPTLLIDHT 442
 DB 7229 RGQVALVL-----GHADATAVHPTRAFDLGFDLSVVELRNRLAAVGLRLPATMVDPY 7284
 OY 443 TPAVVAARLRTAAAGHLDDEAPVPDSSGSGGTAAADDPALITIGMCRFGSGVRSFKDL 502
 DB 7285 TVEVLVDFILDELGG--TEVEAEVVO-----RGTAAVADDPVIVGMCRFGVSGASBDL 7338
 OY 503 WEIASAAGDAIGPPTDGRGWPTGEORHAQDPTQPGTFYPOGGGFLDAAHFDAGFFGISPR 562

DB 7339 WRLVTEGTDVAVSAFVVRGWDVENLXHPDNDNIGTSYTRSGGLHEAGEEDPFGFMSPR 7398
 OY 563 EALAMPQOQLLETSGNEAFERAGIDPLSRGSRGTGVPFAGALSFPDGPMDTSSSEAD 622
 DB 7399 EALSTSQOQLLETSGNEAFERAGIDPLSRGSRGTGVPFAGALSFPDGPMDTSSSEAD 7452
 OY 623 VEGHILGTGVSILSGRIASFGLEGPATVDPGCSALVTLLHACQSLRSGECTLLAG 682
 DB 7453 FEGFGSGSSPSVASGRVSTTFPEFEPAYVVDTRACSSSLVAMMAMQALRRSGEISLALG 7512
 OY 683 GVSVMSTLGMFIERSRQGLSVDRCKAYSAADGTGMEGVMILLVERLSDAVRLGHRV 742
 DB 7513 GVTVMSTPGVFVDFARQGRGLSPGRCKAFESDADGVMSGEVGMVLVERQSDAIRNGHEI 7572
 OY 743 LAVVRSANVODASNGLTAPNGPQERVTROLANAGLSVADVDVEGHTGTLLGDP 802
 DB 7573 LAVVRSANVODASNGLTAPNGPQERVTROLANAGLSVADVDVEGHTGTLLGDP 7632
 OY 803 EAQALLATYGO--RAGDRPLMLGSLKSNIGHTMAAGGVYIKVMALREGLVRLTHVDK 861
 DB 7633 EAQALLATYGRDPEQPLMLGSKNIGHTMAAGGVYIKVMALREGLVRLTHVDK 7692
 OY 862 PSQVDSAGAVRLTEAVPWPQDAAGRLRAGVSSFGIGTNAHVILIE--APAAGCVA 920
 DB 7693 PSQVDSAGAVRLTEAVPWPQDAAGRLRAGVSSFGIGTNAHVILIE--APAAGCVA 7750
 OY 921 GGVLEBAPALASVSAVAPVAVSAPVAPVAPVAPVAPVAPVAPVAPVAPVAPVAPVAPV 980
 DB 7751 GGVLEBAPALASVSAVAPVAVSAPVAPVAPVAPVAPVAPVAPVAPVAPVAPVAPVAPV 7784
 OY 981 RQVAVAPDVSLADVAGLACRAVLEHRAVLAADREELVQGLAAGEPDRVYTTG 1040
 DB 7785 LASVEAELDRPLVDVGSLSVAASLFEHRAVLAADREELVQGLAAGEPDRVYTTG 7844
 OY 1041 APGDRGVVFPVPGGQGAAGVRLASSPVFARMQACEBALPVDVSVVDIL--RR 1099
 DB 7845 APGDRGVVFPVPGGQGAAGVRLASSPVFARMQACEBALPVDVSVVDIL--RR 7901
 OY 1100 DAG---DAVVERADVQVPLFSVMVGLAALMRSYGLBPDAVGHSGOETLAAVCGSLT 1156
 DB 7902 DAG---DAVVERADVQVPLFSVMVGLAALMRSYGLBPDAVGHSGOETLAAVCGSLT 7956
 OY 1157 KDAKTVALRSRALAVRGRGASVPLPAQVEQLIGERWAGRLWAAVNGRSTAVSG 1216
 DB 7957 KDAKTVALRSRALAVRGRGASVPLPAQVEQLIGERWAGRLWAAVNGRSTAVSG 8012
 OY 1217 DAAVDEVLAVCAGTVRARIPVDYASHCPVQPLREELLEGDISPOSGVPFSTV 1276
 DB 8013 DAAVDEVLAVCAGTVRARIPVDYASHCPVQPLREELLEGDISPOSGVPFSTV 8072
 OY 1277 EGTWLDITTL--DAVYVRNLHQVRPSDAVQALADGHRFVSVSPHTLVPALIEDTDE 1335
 DB 8073 EGTWLDITTL--DAVYVRNLHQVRPSDAVQALADGHRFVSVSPHTLVPALIEDTDE 8130
 OY 1336 TAEDVTAIGSLRRGDDTRRFLTALAHHTTGI-----GTPTHHHYTHHTHPR 1386
 DB 8131 TAEDVTAIGSLRRGDDTRRFLTALAHHTTGI-----GTPTHHHYTHHTHPR 8176
 OY 1387 PHTLDELPTVPFOHNYWLESSQPGAGSGAGAGS-----GAGSGRAGTA-- 1432
 DB 8177 PHTLDELPTVPFOHNYWLESSQPGAGSGAGAGS-----GAGSGRAGTA-- 8233
 OY 1433 ---RRVELPTAFQHQFW--PAAPLGGGDVYAAAGLSAHEHLLGAAYLAAGEVLFGR 1461
 DB 8234 ---RRVELPTAFQHQFW--PAAPLGGGDVYAAAGLSAHEHLLGAAYLAAGEVLFGR 8293
 OY 1462 ALHSHPLADHTVDGTVLLPGTALLBLAIRAGEVSDOVELITTAIPLPERGAVQV 1484
 DB 8294 ALHSHPLADHTVDGTVLLPGTALLBLAIRAGEVSDOVELITTAIPLPERGAVQV 8353
 OY 1485 RINTWTVQO-----TWKPL--TLPTTHQPHQTL--IAIPETQTH-----HPHITN- 1526
 DB 8353 RINTWTVQO-----TWKPL--TLPTTHQPHQTL--IAIPETQTH-----HPHITN- 1526

Db 8354 VDAAGCYEEFAAAGFCGRLFOGLRAAMRGDEWFAEVALPEGTEHTAAGFGLHAFDLS 8413
 Qy 1527 -----ILT-----NLHHGITEPIPLTANTHTNPOHLLHHTLH----- 1559
 Db 8414 ALHALLITPADGGTAGLFWSTWVSLHAGATVLRRLTPTGRTLSLAVDSGTPV 8473
 Qy 1560 -----TROOQONHTG-----ALTGLLSLALDETPPHNPHHT 1592
 Db 8474 SVGLSVTRVVDGLTGTTEALLRDAFGEVWTPAQGSEAPATVALICTD-----GD 8525
 Qy 1593 PTGFLMLTLTQHTHTQHTPTPLWYAT-----NATTHNDPLTHPTQ----- 1636
 Db 8526 PLGLAALVTGAGTTRVTRVDAVALVAATVAADGVEPVVNVGASHGSEPSDAHRLTA 8585
 Qy 1637 -----AQTGLARTLLEHPTHTAGIIDL 1661
 Db 8586 DALDTVOSRLSEDRLAGARLVPTRGAVTRDRLGAAVQGLVRSAGTENP- GALGLVLD 8644
 Qy 1662 TPTPTHTLQHTLQTLTPHNOTOLAIRTGTHTRRLPTTLTPTHOPTP- TPHGTLIT 1720
 Db 8645 GTPESAAALAAALGSAEP-----QALRDGRPHARLARLARPADTAAPTAQSDGTIVLT 8700
 Qy 1721 GGTGALALHTLHTLHTTPOHTHLLTSTGTHPTHAQHTLTOLOOKGHLITTCDSNP 1780
 Db 8701 GGTGGLAAGFARHLVDAVGNLILVSRGGDAGTTELVAELTAHGAETVQAADVADP 8760
 Qy 1781 DOLQOLNTTTPROHPLTVITHTAGILDDATLTNLTPTOLNVLRPAKASHLHQLQHT 1840
 Db 8761 DAVTALVAGIPEHPLTAVHTTAGVLDGVTGSLTGERLARVLPKADAMNLLQATRL 8820
 Qy 1841 PLTAFVLYSSAAATFGAPGOANVAAVYDALAHHRHTHLLPATSIAMGTW- OGAGLAD 1899
 Db 8821 DLDAFVSSVAVGFGAGQANVAAAGNAFLDALMEHRAAGLPGLSLAMGPMQDAGMTE 8880
 Qy 1900 --SFKAAVYLDREGRPMSEPLATAVTOALA-----DTEPRVYIADIDMSKI 1946
 Db 8881 TLSAEEARLARACMPRLAAEGFALPDALAAGNGTGTSEDYTGALVVPVRLDAL 8940
 Qy 1947 EHTSOTDLSAAREBPAYORPTPRPAELHKTLA-----HOTSADQRAALL 1992
 Db 8941 -----AAGGEPRVAVRGVTRTRTAVAGTITADGLVARTALTAEERREAL 8989
 Qy 1993 ELVBDHVAALRHADPKAIAPDQSFALGPDSTLAVEFRNLLIKATGLRLEFVSLVDPHT 2052
 Db 8990 GLVTRQALVYGHADAASVDAAQFRDVGFDSTLAVELRNLSATGRLPATLVFDPYN 9049
 Qy 2053 PAKLAVHLQNOJRGTAESAPSAAVTAEASVTERPIAVGACRPPGCVTSADPMDLIS 2112
 Db 9050 AATLASHLHDELFGAETESALALPVKOLPATTDPIVIVGMACRYPGVVASPEDLMRL 9109
 Qy 2113 SEODAIQGFPTDRGMDLDTLYDPDHPGTCYTRNGGFLYDAGHFDABFPGISPREALM 2172
 Db 9110 EGTALITEFPVNRKMDLESVTHPDEHLGTSYTRSGFLHEAGFDEGFFMSREALAT 9169
 Qy 2173 DPOORLLLETAMETIEHAGINPHTLHGTPTGVTGTNGODYALRVHNAAGOSTDFALGT 2232
 Db 9170 DSGORLLLESWEALIERAGIDPVGIRGSAATGVFAGVWVSDYGSIL--GKKEFEGLOQGS 9227
 Qy 2233 AGSIVSGRISTFEGFBEPBPAVBDTACSSSLVALHLCOLAPAGCSMALAGCVTVMSSPG 2292
 Db 9228 AGSVASGRVSYTFEGFBEPBPAVBDTACSSSLVALHLCOLAPAGCSMALAGCVTVMSSPG 9287
 Qy 2293 AFVFSRORGLAAGHCKAPFAAADGTGMEGVGMLVERLSDAHNRGHRVLAIVRGSV 2352
 Db 9288 TTFVFSRORGLSPGRCKAPFSDSADGVMSSEGVGMLVLEROSDAIRNGHEILAVLRSAI 9347
 Qy 2353 NODGASNGLTAPNGPSQORVIRQALANAGLSAGDVDAEBAHGTITGLDPIEAQALLATY 2412
 Db 9348 NODASNGLTAPNGPSQORVIRQALANAGLSAGDVDAEBAHGTITGLDPIEAQALLATY 9407
 Qy 2413 GODDAGCEPLMLGVSXVNGHTOAAAGVAVIYKVMALRHGLRRLTHVDEPSRHYDMSA 2472
 Db 9408 GRDBPEOPLLLGVSXVNGHTOAAAGVAVIYKVMALRHGLRRLTHVDEPSRHYDMSA 9467

Qy 2473 GAVOLLTEVPMPEGBEGRRLRACVSSFGVSGTNAHVLBEAPADVDVPGGPAGEGASD 2532
 Db 9468 GAVELLTECEWETE- HVRACVSSFGISGTNAHYLBE- PAKVIGQTVV-----ASAA 9521
 Qy 2533 DEAAAGPGVPMPLVSAKSOPALRAQALHNLTHPEGLDLDVGTTLAHAPVDPHRA 2592
 Db 9522 PESDVPEPSVPMVLSAKTPEALRDOARLLAAVTEEPARPAPVDVHSLVTRSLFPHRA 9581
 Qy 2593 TLIAADRDFTLOALQALAGEPHRAVHSGAPGCTGGEAGKTAFTCSGQORPCMAH 2652
 Db 9582 VVLAGEHOETVRALTVLAGATDPTAV-----SGT-VASGSALFSGQSORLQMKR 9633
 Qy 2653 GLVHTHPVFAALNDICTHLDPHLDPLRLPTLQONDNDEDAALLOOTRYAAPALFAQ 2712
 Db 9634 ELVGRFPVFAEVLDAVLAVALDGLSGLSREVMGDE-----AGLNETGWTQALFAVE 9687
 Qy 2713 VALHRLITDGYHITPHYVAGHSLGEITAAHLAIGILTDTATTLTORATLMQMP- GTM 2771
 Db 9688 VALYRLV-ESMGVRPDPVAGHSIGETIAAHIAAGVFLSDAARLVAAARGRLMOALPAGGM 9746
 Qy 2772 TLTHTPHHTLHTLHENDLATAINTPSLISGTPHNVCHITTLCOOQIKTKLPT 2831
 Db 9747 VAOATEDEVVPLS--DEVSTALNGPVSVLGVSABVLELARFEAEGRKTRLRV 9803
 Qy 2832 NHAHSPHTNPILOHQTOTLTYHNPHTLTANT-----PPDOLLTPHYWTQOARNTV 2887
 Db 9804 SHAHSPHTNPILOHQTOTLTYHNPHTLTANT-----PPDOLLTPHYWTQOARNTV 2887
 Qy 2888 DVATTTQTLHOGVTVYIELGPDNTLTTLTNHLPNRPPTTLTLTHPHNHTQTLTLNLA 2947
 Db 9864 RFADGVAVLEAQQVTTELFGPDGVSAMAOESLTGDEVTIPLRKDRDESTALLARA 9923
 Qy 2948 KTTT-----WHNNHTHNOPTHT-NDLSTYRQNHNVYLESQFAGVNSAAGLD 3001
 Db 9924 GLYTRIGGGM-----QDLAGTARRNRLTYAFQHOEWPTGRAAAAGVRAAGLG 9976
 Qy 3002 PTEHPRLGATLELATDGCALLAGRLSLRSHPMADAAVAGTIVLSGATLELHAGTVV 3061
 Db 9977 SAHPRLGATLELATDGCALLAGRLSLRSHPMADAAVAGTIVLSGATLELHAGTVV 3061
 Qy 3062 GCDRVDELTHAPLVVPVDSGVSVGVAAADGEGRLVSVVARGSGACGGAGSGVWT 3121
 Db 10037 GCDRVDELTHAPLVVPVDSGVSVGVAAADGEGRLVSVVARGSGACGGAGSGVWT 3121
 Qy 3122 CHASGVLEBAAGVVVDGL-AGVWPBGAVVAVDVGVRRLAGAGCVLGPVSSGLRAYV 3180
 Db 10091 QHAAVYL--APGAGAPETGLGTEWPPRDESLDTGDCYQAFABAGFAYGRLFOGLRAYV 10148
 Qy 3181 RDGGDLAEVCLPEEAMGDAAGFGLHPLALDGVVQPLSVLPPGTFGEGAGGEGVRVP 3240
 Db 10149 RREBELFAEVALLEBEGGEAERFGLHPLALDSA--LHATMLAAATGEBEDG---GGAALP 10202
 Qy 3241 AVWGVSLEHRAVTVGRVRSVAVRGSGREAVSVVVGDEAGVAVSVDRLELBPVMDGL 3300
 Db 10203 FSWEGVALHSGASALRVRLAP---AGDNALSVMHADVAGARVVTGRLTLTGTVDOL 10258
 Qy 3301 RAYSVSNAGR-RGSLVAVQMAEVEGPRVYCGAAMHEDVSGSGGP----- 3344
 Db 10259 KA---AGEQNTLFGDW-----TPVADTRARPASVALVAGADPGLADAFAGAGLTAV 10309
 Qy 3345 -----VPGVVILRCPDAGAGGGGGGGGGGGGVGVGVGVGVGVGVGVGVGVGVGV 3389
 Db 10310 TAADLTALTLASSPAAPDVLV--PAGA-PADGADLPKAVREVTGRVLLALLOOMPVBEA 10366
 Qy 3390 FAGSRLVVTTRGAVVAGPEDEPVVVGAVYGLVRSQAQAEHDPFVLLDLDTGTDLD 3449
 Db 10367 LAASRLVFTVRGA--VAVDDGTAADLAASAVYGLVRSQAQAEHDPFVLLDLDTGTDLD 10419
 Qy 3450 GAGAGKVDGGRVAAVVAOC--EPOLAVGERLLARLXKLESSGDPVAPARSDGTARRS 3507
 Db 10420 -----AVPSARLVEALTAGAEPOLALRGITEVVAARLARLYADAAOPEPA-----S 10463

```
QY 3508 DVPAQRSGVGVARRSVDSGREVLPMWSSGSLVTGTGVLGAAVARRHLAGVCVRDILL 3567
Db 10464 GAMADGTSGT-----FRITGVIVLTGGTGGLGRLLARHLVTHGVRSLIL 10508
QY 3568 VSRGPPDAPGABGRLPAELALGAEVRIYACVGERREYVRLLEGVPAGCPLTGVVHAAGV 3627
Db 10509 ASRRGLQADGAGKLVAELEEMGAEVTVACDVADPEAVAALVAAAPKGRPLSAVVTAGV 10568
QY 3628 LDDATTIASLTPERIGTVFAAKVDAALLDELTRGMELSAFVLFSAGGILGSAGOGNYAA 3687
Db 10569 LDDGVVGSLLTPERLDTVIRPKADAVMHLHRAIRGLDADFVYSSVAGVFGAGQANVAA 10628
QY 3688 ANAALDIALVRRRAAGLPGVSIANGWEASGMTGHLAGTDHRIIRSGLHPMSTPDALA 3747
Db 10629 ANTFLDLIMARRRADGELPGVSIANGPWEQDSGMTGLTDAEERLARSMPPLVVEQHA 10688
QY 3748 LFDALALDRPVLI.P-----ADLRPAPPLPPLQDILLPATRRRTTTRTTGGADNGAQLH 3802
Db 10689 LFDVAVAGERALALPVRIDLAVLRAGEVPALLRGLIRVPARRAAQSAVPSADT--LT 10745
QY 3803 ARLAGQTHEOQHTTLIALVRESHTATVIGHTTPDTIIPPRAPFDLGFDSLTAVELRNLSR 3862
Db 10746 RQLAGLGABQDEEALTLVRQAAITVLGHRDGSALIGREPOFDELGFDSLIAVEFRNRLNT 10805
QY 3863 TTGLRLPTTLAFDHPNPTTLTHHLATQLOPOPDNAVAVLAELDKLESALSALDKTDSAS 3922
Db 10806 ATGLRLPATLTFDYPTPADVVRVLRGQLVTDPSGAGSVLAALGLEKAIATLSDDAEE 10865
QY 3923 -ERYTLRLKSLMLMNNAPQHTASADD--EKFTSATEAEIRFKFIDNDIGLS 3972
Db 10866 HRRVAGRIEVLRTIWAALGR--TDTADEDASSDDLGDASDEDMFALLDDELGLS 10917
```

Search completed: June 17, 2003, 13:07:50
Job time : 304.577 secs